



nestor

nestor-Kriterien Kriterienkatalog vertrauenswürdige digitale Langzeitarchive

Version 2

herausgegeben von der nestor-Arbeitsgruppe
Vertrauenswürdige Archive - Zertifizierung

nestor-Materialien 8



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



nestor-Kriterien
Kriterienkatalog
vertrauenswürdige digitale
Langzeitarchive
- Version 2 -

herausgegeben von der
nestor-Arbeitsgruppe
Vertrauenswürdige Archive -
Zertifizierung

Frankfurt am Main, November 2008

nestor-materialien 8

Impressum

nestor - Materialien 8: nestor - Kompetenznetzwerk Langzeitarchivierung / Arbeitsgruppe
Vertrauenswürdige Archive – Zertifizierung: nestor-Kriterien, Kriterienkatalog
vertrauenswürdige digitale Langzeitarchive, Version 2, 2008, Frankfurt am Main : nestor
c/o Deutsche Nationalbibliothek
urn:nbn:de:0008-2008021802

nestor-Arbeitsgruppe „Vertrauenswürdige Archive - Zertifizierung“

c/o Humboldt-Universität zu Berlin, Universitätsbibliothek
Susanne Dobratz
Unter den Linden 6
D-10099 Berlin
Tel.: +49-30-2093-7070
Fax.: +49-30-2093-2959
E-Mail: dobratz@cms.hu-berlin.de

oder

c/o Bayerische Staatsbibliothek, Digitale Bibliothek
Dr. Astrid Schoger
80328 München
Tel.: +49-89-28638-2600
Fax.: +49-89-28638-2672
E-Mail: astrid.schoger@bsb-muenchen.de

nestor – Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen für Deutschland

c/o Deutsche Nationalbibliothek
Reinhard Altenhöner
Adickesallee 1
D-60322 Frankfurt am Main
E-Mail: lza-info@langzeitarchivierung.de
Web: <http://www.langzeitarchivierung.de>

Autoren/innen des Kriterienkatalogs:

Bergmeyer, Winfried: Institut für Museumsforschung Berlin
Dobratz, Susanne: Humboldt-Universität zu Berlin, Universitätsbibliothek
Dr. Hänger, Andrea: Bundesarchiv Koblenz
Huth, Karsten: Bundesarchiv Koblenz
Dr. Keitel, Christian: Landesarchiv Baden-Württemberg
Dr. Klump, Jens: GeoForschungszentrum Potsdam
Dr. Korb, Nikola: Deutsche Nationalbibliothek Frankfurt
Rödig, Peter: Institut für Softwaretechnologie, Universität der Bundeswehr München
Dr. Rohde-Enslin, Stefan: Institut für Museumsforschung Berlin
Dr. Schoger, Astrid: Bayerische Staatsbibliothek München
Schröder, Kathrin: Bundesarchiv Koblenz
Steinke, Tobias: Deutsche Nationalbibliothek Frankfurt
Strathmann, Stefan: Niedersächsische Staats- und Universitätsbibliothek Göttingen
Prof. Wiesenmüller, Heidrun: Hochschule der Medien Stuttgart

Weitere Anregungen verdanken wir:

Dr. Beckschulte, Klaus: Börsenverein des deutschen Buchhandels, Landesverband Bayern
Kaiser, Max: Österreichische Nationalbibliothek Wien
Dr. Lupprian, Karl-Ernst: Generaldirektion der Staatlichen Archive Bayerns
Dr. Schomburg, Silke: Hochschulbibliothekszentrum Köln
Ullrich, Dagmar: Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen

Beiträge aus der internationalen Diskussion verdanken wir:

Dale, Robin: OCLC/RLG / University of California Santa Cruz (USA)
McHugh, Andrew: Digital Curation Centre (UK)
Reilly, Bernard: Center for Research Libraries (USA)
Prof. Ross, Seamus: Digital Curation Centre (UK)
Ruusalepp, Raivo: Digital Preservation Europe
Waltz, Marie: Center for Research Libraries (USA)

Inhaltsverzeichnis

ZUSAMMENFASSUNG	1
I. Einleitung	3
Grundkonzepte der Langzeitarchivierung digitaler Objekte	3
Der nestor-Kriterienkatalog	7
Grundprinzipien bei der Herleitung von Kriterien	7
Grundprinzipien bei der Anwendung der Kriterien	8
Die nestor-Arbeitsgruppe „Vertrauenswürdige Archive – Zertifizierung“	9
Aktuelle internationale Ansätze zur Evaluierung vertrauenswürdiger digitaler Langzeitarchive	9
II. Kriterienkatalog	11
A. Organisatorischer Rahmen	11
B. Umgang mit Objekten	21
C. Infrastruktur und Sicherheit	39
III. Checkliste	41
IV. Glossar und Abkürzungen	44
V. Bibliographie	47
Anhang	55
Core Requirements for Digital Archives / Kernanforderungen an vertrauenswürdige digitale Langzeitarchive	55

ZUSAMMENFASSUNG

Digitale Informationen sind zu einem nicht mehr wegzudenkenden Bestandteil unseres kulturellen und wissenschaftlichen Erbes geworden. Wissenschaftliche Erkenntnisse, historische Dokumente und kulturelle Leistungen liegen in stark zunehmendem Maße – und häufig ausschließlich – in elektronischer Form vor. Neben den unschätzbaren Vorteilen dieser Form besteht ein gravierender Nachteil: Zur Erfassung der Information durch den Menschen ist erheblicher technischer Aufwand unabdingbar. Zusätzlich unterliegen die benötigten technischen Mittel nach wie vor einer unvorstellbar rasanten Weiterentwicklung. Die damit verbundene schnelle Überalterung der interpretierenden Technik ist in Kombination mit dem oftmals unsichtbaren physischen Verfall der Medien eine ernste Bedrohung für den Informationserhalt – nicht nur langfristig!

Dieser Umstand wirft die begründete Frage nach Vertrauenswürdigkeit auf. Produzenten und Konsumenten von Information stellen sich die Frage, welche Gedächtnisorganisationen überhaupt in der Lage sind, Authentizität, Integrität, Vertraulichkeit und Verfügbarkeit digitaler Information zu wahren. Und Verantwortliche in Institutionen, konfrontiert mit einer unaufhaltsamen Flut von digitalen Objekten, stehen vor der Herausforderung, Vertrauenswürdigkeit zu erlangen und nach außen darzustellen, sei es um der Erfüllung ihres gesetzlichen Auftrags gerecht zu werden oder um auf dem Markt zu bestehen.

Genau hier setzt die Arbeit der nestor-Arbeitsgruppe „Vertrauenswürdige Archive – Zertifizierung“ an. Sie identifiziert Kriterien, die die Bewertung der Vertrauenswürdigkeit eines digitalen Langzeitarchivs sowohl in organisatorischer als auch in technischer Hinsicht ermöglichen. Die Definition der Kriterien geschieht in engem Kontakt mit unterschiedlichsten Gedächtnisorganisationen und Produzenten von Information sowie weiteren Betroffenen und Experten. Dieses offene Vorgehen garantiert eine hohe Allgemeingültigkeit und Praxistauglichkeit sowie eine breite Akzeptanz der Ergebnisse. Der nun vorliegende Kriterienkatalog bildet einen wichtigen Meilenstein beim Erreichen der Ziele der Arbeitsgruppe. Einerseits soll den Gedächtnisorganisationen damit ein fundiertes, abgestimmtes und praxisiertes Hilfsmittel zur Erlangung und Darstellung von Vertrauenswürdigkeit bereitgestellt werden. Andererseits soll auch die Option eröffnet werden, die Vertrauenswürdigkeit durch eine Zertifizierung im Rahmen eines national bzw. international standardisierten Verfahrens nachzuweisen. Der Katalog unterstützt außerdem die aktive Teilnahme an bereits laufenden internationalen Standardisierungsarbeiten.

Um einen breit akzeptierten Kriterienkatalog entwickeln zu können, hat sich nestor der Beiträge und Kommentare betroffener und interessierter Institutionen bedient. Die Arbeitsgruppe hat von Anfang an ein offenes Verfahren gewählt, um an der Problematik zu arbeiten und frühzeitig alle Interessengruppen einzubeziehen. Die vorliegende

Fassung berücksichtigt auch die aktuellen Diskussionen unserer internationalen Partner insbesondere der OCLC/RLG–NARA Digital Repository Certification Task Force, dem Center for Research Libraries, dem Digital Curation Centre und dem Projekt Digital Preservation Europe.

Das vorliegende Dokument gibt zunächst eine kurze Einführung in die Problematik der Langzeitarchivierung digitaler Objekte. Eine Beschreibung wichtiger Konzepte und Prinzipien, die dem Kriterienkatalog zu Grunde liegen, sorgt für eine leichte Verständlichkeit. Die Ziele und das Vorgehen der Arbeitsgruppe werden kurz erläutert. Dieser Einleitung folgt der Kriterienkatalog in seiner ausführlichen Form. Eine kompakte Darstellung des Katalogs als Checkliste und ein Glossar runden das vorliegende Dokument ab.

I. Einleitung

Grundkonzepte der Langzeitarchivierung digitaler Objekte

Bedrohungen für den Informationserhalt, Vertrauenswürdigkeit

Informationen, die durch digitale Objekte repräsentiert sind, sind bedroht durch Einbußen in ihrer Integrität, Authentizität und Vertraulichkeit sowie den gänzlichen Verlust der Verfügbarkeit und Nutzbarkeit. Besondere Herausforderungen für die Langzeitarchivierung stellen die nicht offensichtliche Bindung der Information an die Datenträger, die physische Alterung der Datenträger sowie die rapiden Veränderungen der für die Interpretation der digitalen Objekte erforderlichen technischen Infrastruktur dar.

Digitale Langzeitarchive haben den Erhalt der Informationen über lange Zeiträume hinweg zum Ziel. Deshalb ergreifen sie sowohl organisatorische als auch technische Maßnahmen, um diesen Bedrohungen entgegenzuwirken. Vertrauenswürdige digitale Langzeitarchive operieren nach ihren Zielen und Spezifikationen. Anhand eines Kriterienkatalogs kann die Vertrauenswürdigkeit digitaler Langzeitarchive geprüft und bewertet werden.

Information, Repräsentation, Interpretation

Ein **Informationsobjekt** ist eine logisch abgegrenzte Informationseinheit. Ein Informationsobjekt kann (teilweise oder vollständig) durch digitale Objekte repräsentiert werden. Dies kann auf verschiedene Weisen, d.h. mit unterschiedlichen Repräsentationen, geschehen (z.B. kann ein literarisches Werk durch eine digitale Volltextpublikation als Textdatei, durch ein Hörbuch als Audiodatei oder durch ein retrospektives Image-Digitalisat einer gedruckten Ausgabe repräsentiert werden). Zur Nutzung der Informationen müssen die digitalen Objekte interpretiert werden, so dass eine maschinelle Weiterverarbeitung oder eine Erfassung durch den Menschen möglich ist. Dies geschieht mit Hilfe so genannter Repräsentationsinformation (z.B. Beschreibung des verwendeten Text-, Bild-, Audioformates, Beschreibung einer Datenbankstruktur etc.). Um die durch digitale Objekte repräsentierten Informationen zu erhalten, muss ein digitales Langzeitarchiv somit die Repräsentationen zusammen mit den entsprechenden Repräsentationsinformationen erhalten.

Der Informationsbegriff umfasst hier jeden Typ von Wissen, das ausgetauscht werden kann, aus inhaltlicher Sicht also etwa Werke geistiger Schöpfung, Ergebnisse der Forschung und Entwicklung, Dokumentationen des politischen, sozialen und wirtschaftlichen Handelns.

Digitale Objekte sind häufig in Form von Dateien organisiert. Ein digitales Objekt kann sowohl aus einer einzigen Datei (z.B. ein digitales Foto als TIFF-Datei) als auch aus mehreren Dateien bestehen (z.B. eine elektronische Zeitschrift mit einzelnen Artikeln als PDF-Dateien). Darüber hinaus kann eine Datei auch mehrere digitale Objekte umfassen (z.B. eine Datenbankdatei).

Inhaltsdaten, Metadaten, Digitale Objekte

Daten repräsentieren Information in einer formalisierten Art, die die Interpretation, die Verarbeitung und den Austausch erlauben. Digitale Objekte sind Verwaltungseinheiten digitaler Daten in einem informationstechnischen System.

Zu den digitalen Daten, die ein Informationsobjekt repräsentieren (Inhaltsdaten), können neben den Daten, die der Interpretation dienen (Repräsentationsmetadaten) noch weitere Daten hinzukommen, die z.B. der Identifizierung, Auffindbarkeit, dem Nachweis der Integrität und Authentizität sowie der Kontrolle der Nutzungsrechte dienen. Diese als Metadaten bezeichneten Daten können zu unterschiedlichen Zeiten im Lebenszyklus digitaler Objekte entstehen (z.B. bei der Produktion, bei der Archivierung, bei der Bereitstellung für die Nutzung). Sie bilden zusammen mit den Inhaltsdaten eine konzeptionelle Einheit – ein Informationspaket. Je nach Bearbeitungsstufe im digitalen Langzeitarchiv werden die Informationspakete eingeteilt in: Übergabepakete, Archivpakete und Nutzungspakete.

Bei einer Implementierung können Metadaten sowohl getrennt (z.B. in einem Datenbanksystem) als auch gemeinsam mit den Inhaltsdaten (z.B. in einer einzigen Datei) gespeichert und verwaltet werden.

Digitales Langzeitarchiv

Unter einem digitalen Langzeitarchiv wird in diesem Kriterienkatalog eine Organisation (bestehend aus Personen und technischen Systemen) verstanden, die die Verantwortung für den Langzeiterhalt und die Langzeitverfügbarkeit digitaler Objekte sowie für ihre Interpretierbarkeit zum Zwecke der Nutzung durch eine bestimmte Zielgruppe übernommen hat. Dabei bedeutet „Langzeit“: über Veränderungen in der Technik (Soft- und Hardware) hinweg und unter Berücksichtigung möglicher Änderungen auch der Zielgruppe. Diese Definition eines digitalen Langzeitarchivs orientiert sich am Referenzmodell für ein „Open Archival Information System“ (OAIS) [1].¹

Das digitale Langzeitarchiv kann Bestandteil einer größeren Einrichtung sein, die auch konventionelle Objekte archiviert. Die Zusammenhänge zwischen analogen und

¹ Der Begriff „Langzeitarchivierung“ wird verwendet, um eine Abgrenzung zu dem Begriff „Archivierung“ der Informationstechnologie vorzunehmen, wo „Archivierung“ durchaus kürzere Zeiträume ohne Veränderungen in der Technik bzw. der Zielgruppe betreffen kann. Z.B. ZIP-Archive, TAR-Archive etc.

digitalen Objekten sollten gewahrt und bei der Recherche entsprechend dargestellt werden.

Nutzung durch Zielgruppe

Der Informationserhalt im Hinblick auf die Nutzung wird als zentrale Aufgabe der Langzeitarchivierung definiert, die in Abhängigkeit von der Zielgruppe und deren Bedürfnissen geleistet werden muss. Nutzung setzt zum einen den Erhalt der Integrität, Authentizität und Vertraulichkeit sowie der Verfügbarkeit der digitalen Objekte voraus, zum anderen die Sicherstellung der Interpretierbarkeit der digitalen Objekte durch die Zielgruppe, um die darin enthaltene Information in einer geeigneten Weise zu rekonstruieren. Änderungen der rechtlichen und organisatorischen Rahmenbedingungen sowie technische Entwicklungen können zu Veränderungen der Zielgruppen und deren Bedürfnisse führen. Ein digitales Langzeitarchiv beobachtet diese Veränderungen und reagiert darauf.

Vertrauenswürdigkeit

Vertrauenswürdigkeit (engl. trustworthiness) wird als Eigenschaft eines Systems angesehen, gemäß seinen Zielen und Spezifikationen zu operieren (d.h. es tut genau das, was es zu tun vorgibt). Aus Sicht der IT-Sicherheit stellen Integrität, Authentizität, Vertraulichkeit und Verfügbarkeit Grundwerte dar. IT-Sicherheit ist somit ein wichtiger Baustein für vertrauenswürdige digitale Langzeitarchive.

Das Spektrum der bestehenden und im Aufbau befindlichen digitalen Langzeitarchive ist sehr groß, wie die folgenden Beispiele zeigen:

Beispiel 1: Eine große wissenschaftliche Universalbibliothek, die die Verantwortung für stetig wachsende Sammlungen digitaler Veröffentlichungen von Verlagen und Ämtern, für wissenschaftlich relevante Internetressourcen sowie für die Ergebnisse von Digitalisierungsprojekten etc. übernommen hat. Die Zielgruppe dieses digitalen Langzeitarchivs ist die Allgemeinheit. Die Produzenten sind vielfältig: Verlage, Ämter, Digitalisierungszentren, verlegende Institutionen, Privatpersonen etc. Eine solche Bibliothek kann die Langzeitarchivierung für kleinere Einrichtungen als Dienstleistung übernehmen; ferner kann sie in ein Netzwerk eingebunden sein, das die Kooperation mit anderen Bibliotheken ermöglicht und dem Nutzer einen einheitlichen Zugriff auf kooperativ aufgebaute Bestände erlaubt.

Beispiel 2: Eine Universitätsbibliothek, die neben kommerzieller wissenschaftlicher Literatur auch E-Learning-Medien, Hochschulschriften, Veröffentlichungen von

Hochschulangehörigen etc. verwaltet. Hier sind die Nutzer Studierende und Mitarbeiter der Universität. Die Produzenten sind zum großen Teil Hochschulangehörige.

Beispiel 3: Eine Forschungseinrichtung, die selbst große Mengen von Fachdaten erzeugt und archiviert. Deren Zielgruppe sind Fachwissenschaftler, die das Spezialwissen zur Interpretation dieser Daten mitbringen.

Beispiel 4: Ein Archiv, das auf der Grundlage von Archivgesetzen die elektronischen Unterlagen von Verwaltungen archiviert. Die Zielgruppe sind neben der Allgemeinheit die Produzenten selbst. Die Nutzung kann aufgrund von Schutzrechten über längere Zeiträume verboten sein.

Beispiel 5: Ein Museum, das Digitalisierungen von Museumsobjekten, aber auch originale digitale Kunst verwaltet. Die Nutzer sind das allgemeine Publikum, Kunstwissenschaftler, Künstler etc.

Beispiel 6: Ein Dienstleister, der im Auftrag anderer Einrichtungen für deren Sammlungen die Langzeitarchivierung übernimmt. Für den Aufbau der Sammlungen sind die Einrichtungen zuständig, der Dienstleister bietet den sicheren Erhalt der digitalen Objekte und deren Verfügbarkeit und Nutzbarkeit an.

Der Weg zum vertrauenswürdigen digitalen Langzeitarchiv

Ein digitales Langzeitarchiv entsteht als komplexer Gesamtzusammenhang. Die Umsetzung der einzelnen Kriterien muss stets vor dem Hintergrund der Ziele des Gesamtsystems gesehen werden. Sowohl die Realisierung des digitalen Langzeitarchivs als Ganzes als auch die Erfüllung der einzelnen Kriterien läuft als Prozess in mehreren Stufen ab:

1. Konzeption
2. Planung und Spezifikation
3. Umsetzung und Implementation
4. Evaluierung

Diese Stufen sind nicht als starres Phasenmodell zu betrachten. Vielmehr müssen sie im Zuge der ständigen Verbesserung regelmäßig wiederholt werden. Das Qualitätsmanagement überwacht diesen Entwicklungsprozess.

Der nestor-Kriterienkatalog

Anwender des Kriterienkatalogs

Der vorliegende Kriterienkatalog richtet sich in erster Linie an Gedächtnisorganisationen (Archive, Bibliotheken, Museen) und dient als Leitfaden, um ein vertrauenswürdiges digitales Langzeitarchiv zu konzipieren, zu planen und umzusetzen. Ferner kann er auf allen Stufen der Entwicklung zur Selbstkontrolle eingesetzt werden.

Darüber hinaus soll dieser Katalog allen Institutionen, die selbst archivieren, Dienstleistern aus dem kommerziellen und nichtkommerziellen Bereich sowie Drittanbietern von Produkten als Orientierungshilfe dienen.

Grundprinzipien bei der Herleitung von Kriterien

Abstraktion

Ziel dieses Kataloges ist es, Kriterien zu formulieren, die für ein breites Spektrum digitaler Langzeitarchive angewendet werden können und über längere Zeit Gültigkeit behalten sollen. Deshalb wird von relativ abstrakten Kriterien ausgegangen. Den Kriterien werden jeweils ausführliche Erläuterungen und konkrete Beispiele aus verschiedenen Bereichen mitgegeben. Die Beispiele entsprechen dem heutigen Stand der Technik und Organisation und sind unter Umständen nur im Kontext einer spezifischen Archivierungsaufgabe sinnvoll. Sie haben keinen Anspruch auf Vollständigkeit.

Anlehnung an die Begrifflichkeit des OAIS

Als Basis für eine gemeinsame Begrifflichkeit und zur Strukturierung des Kriterienkataloges dient – wo möglich – das Referenz-Modell OAIS mit seinen funktionalen Entitäten und dem Informationsmodell. Einerseits wird das OAIS zur Beschreibung der Kernprozesse von der Aufnahme der digitalen Objekte ins digitale Langzeitarchiv (Ingest) über die Archivablage (Archival Storage) hin zur Nutzung (Access), andererseits zur Beschreibung des Lebenszyklus digitaler Objekte vom Produzenten über das digitale Langzeitarchiv hin zum Nutzer herangezogen. Dabei werden folgende Informationspakete berücksichtigt: Übergabepaket (Submission Information Package, SIP) bei der Aufnahme, Archivpaket (Archival Information Package, AIP) in der Archivablage und Nutzungspaket (Dissemination Information Package, DIP) bei der Nutzung.

Grundprinzipien bei der Anwendung der Kriterien

Dokumentation

Die Ziele, die Konzeption und Spezifikation sowie die Implementierung des digitalen Langzeitarchivs sind angemessen zu dokumentieren. Dies erlaubt zu überprüfen, ob alle Teile des digitalen Langzeitarchivs aufeinander abgestimmt sind. Anhand der Dokumentation kann der Entwicklungsstand intern und extern bewertet werden. Eine frühzeitige Bewertung kann dazu dienen, Fehler durch eine ungeeignete Implementierung zu vermeiden. Alle Qualitäts- und Sicherheitsnormen fordern eine angemessene Dokumentation.

Transparenz

Transparenz wird realisiert durch die Veröffentlichung geeigneter Teile der Dokumentation. Transparenz nach außen gegenüber Nutzern und Partnern ermöglicht diesen, selbst den Grad an Vertrauenswürdigkeit festzustellen. Transparenz gegenüber Produzenten und Lieferanten bietet diesen die Möglichkeit zu bewerten, wem sie ihre digitalen Objekte anvertrauen.

Transparenz nach innen dokumentiert gegenüber den Betreibern, den Trägern, dem Management sowie den Mitarbeitern die angemessene Qualität des digitalen Langzeitarchivs und sichert die Nachvollziehbarkeit der Maßnahmen.

Bei denjenigen Teilen der Dokumentation, die für die breite Öffentlichkeit nicht geeignet sind (z.B. Firmengeheimnisse, Informationen mit Sicherheitsbezug), kann die Transparenz auf einen ausgewählten Kreis (z.B. zertifizierende Stelle) beschränkt werden.

Durch das Prinzip der Transparenz wird Vertrauen aufgebaut, da es die unmittelbare Bewertung der Qualität eines digitalen Langzeitarchivs durch Interessierte zulässt.

Angemessenheit

Das Prinzip der Angemessenheit berücksichtigt die Tatsache, dass keine absoluten Maßstäbe möglich sind, sondern dass sich die Bewertung immer an den Zielen und Aufgaben des jeweiligen digitalen Langzeitarchivs ausrichtet. Die Kriterien müssen im Kontext der jeweiligen Archivierungsaufgabe gesehen werden. Deshalb können ggf. einzelne Kriterien irrelevant sein. Auch der notwendige Erfüllungsgrad eines Kriteriums kann – je nach den Zielen und Aufgaben des digitalen Langzeitarchivs – unterschiedlich ausfallen.

Bewertbarkeit

Für die Vertrauenswürdigkeit existieren zum Teil – insbesondere unter Langzeitaspekten – keine objektiv bewertbaren (messbaren) Merkmale. In diesen Fällen weisen Indikatoren, wie etwa das Vorhandensein angemessener Metadaten, auf den

Grad der Vertrauenswürdigkeit hin. Transparenz macht auch die Indikatoren für eine Bewertung zugänglich.

Die nestor-Arbeitsgruppe „Vertrauenswürdige Archive – Zertifizierung“

Um in Deutschland zu einem ersten Katalog von Kriterien der Vertrauenswürdigkeit zu kommen und die Zertifizierung von digitalen Langzeitarchiven nach national und international abgestimmten Verfahren vorzubereiten, wurde im Rahmen des vom BMBF geförderten Projekts nestor die Arbeitsgruppe „Vertrauenswürdige Archive – Zertifizierung“ gegründet. Die Mitglieder der Arbeitsgruppe sind Vertreter aus Bibliotheken, Archiven, Museen, Forschungseinrichtungen, Verlagen sowie Software- und Zertifizierungsexperten.

Der vorliegende Katalog wurde in einem offenen Prozess unter Einbeziehung weiterer Experten und Einrichtungen mit praktischer Erfahrung im Aufbau digitaler Langzeitarchive entwickelt. Als Grundlage dienten die Ergebnisse einer Umfrage und eines offenen Workshops. Der erste Entwurf des Kriterienkataloges wurde in einem Expertengespräch vorgestellt und diskutiert. Im Juni 2006 wurde der Katalog als Entwurf zur öffentlichen Kommentierung publiziert. Die vielfachen Rückmeldungen sind in Version 2 ebenfalls eingeflossen.

Um den nestor-Kriterienkatalog in die internationale Diskussion einzubringen wurde Version 1 im Dezember 2006 auch in englischer Sprache veröffentlicht.

Der nestor-Katalog fokussiert auf die Anwendung in Deutschland, er wird jedoch auch im internationalen Kontext diskutiert und standardisiert werden. Dafür ist es entscheidend, allgemeingültige Kriterien vor dem Hintergrund der spezifisch nationalen Rahmenbedingungen zu identifizieren. Diese liegen u.a. in den rechtlichen Voraussetzungen, der Ausstattung öffentlicher Einrichtungen mit finanziellen und personellen Ressourcen, in nationalen organisatorischen Strukturen und dem nationalen Entwicklungsstand im Bereich der digitalen Langzeitarchivierung begründet.

Aktuelle internationale Ansätze zur Evaluierung vertrauenswürdiger digitaler Langzeitarchive

Der nestor-Kriterienkatalog berücksichtigt nationale und internationale Ansätze und Arbeitsergebnisse wie das DINI-Zertifikat für Dokumenten- und Publikationsserver [2], den RLG-OCLC-Bericht „Trusted Digital Repositories: Attributes and Responsibilities“ (Mai 2002) [3] und den von der OCLC/RLG-NARA Digital Repository Certification Task Force veröffentlichten Entwurf von August 2005 und Endfassung von Februar 2007 unter dem Titel „Trustworthy Repositories Audit & Certification: Criteria and Checklist“ (TRAC) [4].

Das Werkzeug *DRAMBORA* [5] wurde vom Digital Curation Centre (DCC) [6] in Kooperation mit dem Projekt Digital Preservation Europe (DPE) [7] entwickelt. Es unterstützt auf der Grundlage der bestehenden Kriterienkataloge eine Analyse eines digitalen Langzeitarchivs von den Zielen über die abgeleiteten Aufgaben zu den bestehenden Risiken und den eingesetzten Gegenmaßnahmen und erlaubt damit die Fähigkeiten, Stärken und Schwächen des Systems aufzudecken.

Zudem kooperierte die Arbeitsgruppe mit der OCLC/RLG-NARA Digital Repository Certification Task Force sowie dem Projekt „Certification of Digital Archives Project“ des Centre for Research Libraries (CRL). Nun arbeitet sie zusammen mit dem „Long-Lived Digital Collections“ Project des CRL [8] sowie dem Digital Curation Centre und dem EU-Projekt Digital Preservation Europe.

Die internationale Zusammenarbeit der obigen Initiativen mündete im Januar 2007 in der Formulierung von 10 Kernanforderungen an vertrauenswürdige digitale Langzeitarchive [9] (eine deutsche Übersetzung findet sich im Anhang).

Aus Gründen der Übersichtlichkeit wird im folgenden Katalog der Term „digitales Langzeitarchiv“ mit „dLZA“ abgekürzt. Ferner wird nur die männliche Bezeichnung verwendet, gemeint sind aber stets Frauen und Männer in diesen Rollen.

Den Aufbau des Kriterienkatalogs zeigt die folgende Übersicht:

Kriterium
Allgemeine Erläuterungen des Kriteriums
Beispiele, Anmerkungen, Hinweise aus den unterschiedlichen Anwendungsbereichen ohne Anspruch auf Vollständigkeit
<i>Literatur, die sich auf dieses Kriterium bezieht</i>

II. Kriterienkatalog

A. Organisatorischer Rahmen

Das digitale Langzeitarchiv agiert in einem organisatorischen Rahmen, der durch die Zieldefinition, die rechtlichen Bedingungen, die personellen und finanziellen Ressourcen gegeben ist.

1 Das digitale Langzeitarchiv hat seine Ziele definiert.

Das dLZA ist sich über seine Ziele im Klaren. Es hat festgelegt, welche Aufgaben es erfüllt und welchen Prinzipien es dabei folgt. Dies ist unabdingbar, weil Vertrauenswürdigkeit kein absoluter Begriff ist, sondern auf die Ziele des jeweiligen dLZA zu beziehen ist. Dem Prinzip der Angemessenheit folgend, richtet sich die Bewertung der einzelnen Kriterien stets an den Zielen aus. Das dLZA macht seine Ziele transparent, damit auch andere – insbesondere Nutzer und Produzenten – die Vertrauenswürdigkeit überprüfen können. (Die Ziele werden häufig in Form einer Policy veröffentlicht.)

PANDORA: The purpose of the PANDORA Archive,
<http://pandora.nla.gov.au/overview.html>

Oxford Digital Library: Background, Services, Principles and Guidelines,
<http://www.odl.ox.ac.uk/about.htm>

Dokumenten- und Publikationsserver der Humboldt-Universität zu Berlin:
Ziele und inhaltliche Kriterien, http://edoc.hu-berlin.de/e_info/leitlinien.php

National Archives: Custodial policy for digital records,
<http://www.nationalarchives.gov.uk/recordsmanagement/custody/>

National Archives and Records Administration (NARA, USA): ERA Vision Statement, <http://www.archives.gov/era/about/vision.html>

Erpanet-Tagung: "Policies for Digital Preservation", 2003. [10]

1.1 Das digitale Langzeitarchiv hat Kriterien für die Auswahl seiner digitalen Objekte entwickelt.

Das dLZA hat festgelegt, auf welche digitalen Objekte sich seine Verantwortlichkeit erstreckt. Häufig leitet sich dies vom Gesamtauftrag der Institution ab oder ist durch Gesetze vorgegeben. Das dLZA hat Sammelrichtlinien, Auswahlkriterien, Bewertungskriterien oder Kriterien der Überlieferungsbildung entwickelt. Die Kriterien können inhaltlicher, formaler oder qualitativer Art sein.

Bei staatlichen und nicht-staatlichen Archiven ist die formale

Zuständigkeit in der Regel durch die gesetzliche Grundlage oder den Archivträger vorgegeben (ein Staatsarchiv übernimmt die Unterlagen der Landesverwaltung, ein Unternehmensarchiv die Unterlagen des Unternehmens, ein Universitätsarchiv die Unterlagen der Universität).

Gesetz über die Deutsche Nationalbibliothek (DNBG,
<http://www.bundesrecht.juris.de/dnbg/BJNR133800006.html>)

§2 Aufgaben und Befugnisse:

Die Bibliothek hat die Aufgabe,

1. a) die ab 1913 in Deutschland veröffentlichten Medienwerke und
- b) die ab 1913 im Ausland veröffentlichten deutschsprachigen Medienwerke, Übersetzungen deutschsprachiger Medienwerke in andere Sprachen und fremdsprachigen Medienwerke über Deutschland im Original zu sammeln, zu inventarisieren, zu erschließen und bibliografisch zu verzeichnen, auf Dauer zu sichern und für die Allgemeinheit nutzbar zu machen sowie zentrale bibliothekarische und nationalbibliografische Dienste zu leisten.

Das von den Landesbibliotheken getragene Baden-Württembergische Online-Archiv (<http://www.boa-bw.de/>) sammelt Netzpublikationen, „die in Baden-Württemberg erscheinen oder sich inhaltlich mit dem Land, seinen Orten und Personen beschäftigen“.

The Oxford Text Archive (<http://ota.ahds.ac.uk/>) sammelt „high-quality scholarly electronic texts and linguistic corpora (and any related resources) of long-term interest and use across the range of humanities disciplines“. Die Website enthält eine detaillierte „Collections Policy“.

Auf dem Dokumenten- und Publikationsserver der Humboldt-Universität zu Berlin werden „durch Mitarbeiterinnen und Mitarbeiter der Humboldt-Universität veröffentlichte elektronische wissenschaftliche Dokumente“ gesammelt (http://edoc.hu-berlin.de/e_info/leitlinien.php).

Erpanet-Tagung: "Appraisal of Scientific Data", 2003. [11]

Wiesenmüller, Heidrun; Jendral, Lars et al.: Auswahlkriterien für das Sammeln von Netzpublikationen im Rahmen des elektronischen Pflichtexemplars: Empfehlungen der Arbeitsgemeinschaft der Regionalbibliotheken, 2004. [12]

1.2 Das digitale Langzeitarchiv übernimmt die Verantwortung für den dauerhaften Erhalt der durch die digitalen Objekte repräsentierten Informationen.

Das dLZA bekennt sich explizit zu seiner Verantwortung für die Langzeitarchivierung der gemäß 1.1 übernommenen digitalen Objekte. Unter Langzeitarchivierung ist dabei der dauerhafte Erhalt der Benutzbarkeit der durch die digitalen Objekte repräsentierten Informationen zu verstehen (vgl. Informationsmodell des OAIS).

Formulierung auf der Website des Internet Archive

(<http://www.archive.org/about/about.php>): „The Internet Archive is working to prevent the Internet (...) and other "born-digital" materials from disappearing into the past. Collaborating with institutions including the Library of Congress and the Smithsonian, we are working to preserve a record for generations to come.“

Formulierung auf der Website der Oxford Digital Library

(<http://www.odl.ox.ac.uk/principles.htm>): „Like traditional collection development long-term sustainability and permanent availability are major goals for the Oxford Digital Library.“

1.3 Das digitale Langzeitarchiv hat seine Zielgruppe(n) definiert.

Zu den allgemeinen Festlegungen für ein dLZA gehört auch die Definition der Zielgruppe(n). Dies beinhaltet die Kenntnis der spezifischen Anforderungen der Zielgruppe(n), die die Auswahl der anzubietenden Dienstleistungen beeinflusst.

Sollten sich die Zielgruppe(n) bzw. deren Anforderungen im Laufe der Zeit ändern, so reagiert das dLZA entsprechend durch Anpassung seiner Dienstleistungen.

Zielgruppen können sein:

- Angehörige einer Behörde, eines Forschungsinstituts etc.
- Fachwissenschaftler einer bestimmten Disziplin
- Allgemeinheit

2 Das digitale Langzeitarchiv ermöglicht seinen Zielgruppe(n) eine angemessene Nutzung der durch die digitalen Objekte repräsentierten Informationen.

Das dLZA sieht seine zentrale Aufgabe in der momentanen und zukünftigen Nutzung der durch die digitalen Objekte repräsentierten Informationen durch seine Zielgruppe(n). Die Nutzung setzt den Erhalt der digitalen Objekte und deren Verfügbarkeit sowie die Sicherstellung der Interpretierbarkeit voraus. Die Nutzung kann angemessen sein, obwohl diese z.B. aus rechtlichen Gründen eingeschränkt wird (vgl. 3.3) oder nicht alle Eigenschaften des Originals erhalten wurden (vgl. 9.2).

So genannte „dark archives“ werden zunächst ohne jede Nutzungsmöglichkeit aufgebaut; sie kommen nur bei etwaigem Ausfall des primären Archivs zum Einsatz. Für einen solchen Krisenfall muss aber auch hier die Nutzung möglich sein.

2.1 Das digitale Langzeitarchiv ermöglicht seinen Zielgruppe(n) den Zugang zu den digitalen Objekten.

Das dLZA stellt sicher, dass berechtigte Nutzer einen Zugang zu den digitalen Objekten erhalten. Dazu gehören angemessene Recherchemöglichkeiten. Bei der Festlegung seines Dienstleistungsportfolios berücksichtigt das dLZA die Bedürfnisse seiner Zielgruppe(n). Das dLZA macht seine Nutzungsbedingungen sowie etwaige für die Nutzung anfallende Kosten vorab bekannt und schlüsselt diese in transparenter Weise auf.

Der Zugang kann ermöglicht werden durch:

- Zugreifen auf die digitalen Objekte
- Erstellen bzw. Lieferung einer analogen Kopie (z.B. als Ausdruck durch den Nutzer oder im Print-on-demand-Verfahren)
- Erstellen bzw. Lieferung einer digitalen Kopie (z.B. Download auf einem Speichermedium des Nutzers, E-Mail-Versand)
- Schnittstellen, um anderen Systemen den Zugriff auf die digitalen

Objekte zu ermöglichen.

2.2 Das digitale Langzeitarchiv stellt die Interpretierbarkeit der digitalen Objekte durch seine Zielgruppe(n) sicher.

Das dLZA trifft geeignete Maßnahmen, um die dauerhafte Interpretierbarkeit der digitalen Objekte zu gewährleisten, und schafft damit die Grundvoraussetzung für eine angemessene Nutzung. Dies umfasst die Interpretierbarkeit der Inhalts- und Metadaten.

Das dLZA berücksichtigt dabei die Bedürfnisse seiner Zielgruppe(n). Je spezialisierter die Zielgruppe(n) sind, desto mehr an Know-how und technischer Ausstattung (z.B. bestimmte Software) bzw. Bereitschaft, zusätzliche Ausstattung einzurichten (z.B. Installieren von Plugins), kann vorausgesetzt werden.

Änderungen der technischen Umgebung oder der Zielgruppe(n) können Einfluss auf die Interpretierbarkeit haben. Das dLZA überprüft deshalb regelmäßig mit geeigneten Verfahren, ob die Interpretierbarkeit durch die Zielgruppe(n) noch gegeben ist.

Maßnahmen können sein:

- Konvertierung in aktuelles Standardformat
- Bereitstellung von Emulationen (z.B. Open Source DOS-Emulator „DOSBox“)
- Bereitstellung der Repräsentationsinformation: z.B. Dokumentation von Datenstrukturen und Feldinhalten, damit Nutzer Daten aus Fachanwendungen (Datenbanken) in jeweils aktuelle Datenbankanwendungen einlesen können
- Bereitstellung von Gebrauchsanweisungen, Installationsanweisungen, Hilfetexten
- Übernahme von Recherchen bzw. Auswertungen durch das dLZA als (kostenpflichtiges) Serviceangebot
- Überprüfen der Interpretierbarkeit anhand regelmäßiger Stichproben
- Angebot eines Feedback-Formulars, mit dem Nutzer Probleme bei der Interpretation melden können

3 Gesetzliche und vertragliche Regelungen werden eingehalten.

Das dLZA agiert auf der Basis rechtlicher Regelungen. Diese betreffen sowohl den Bereich der Übernahme der digitalen Objekte als auch die Archivierung und Nutzung. Dabei schafft das dLZA einen Ausgleich zwischen den berechtigten Interessen der Produzenten und denen der Nutzer sowie ggf. der Betroffenen (im Falle personenbezogener Daten).

BArchG: Gesetz über die Sicherung und Nutzung von Archivgut des Bundes, 2005. [13]

Goebel, Jürgen W.; Scheller, Jürgen et al.: nestor - materialien 1: Digitale Langzeitarchivierung und Recht, 2004. [14]

Coyle, Karen: Rights in the PREMIS Data Model. A Report for the Library

of Congress, 2006. [15]

3.1 Es bestehen rechtliche Regelungen zwischen Produzenten und dem digitalen Langzeitarchiv.

Um Planungs- und Rechtssicherheit zu schaffen, trifft das dLZA – soweit möglich – formale Übereinkünfte mit den Produzenten bzw. Lieferanten. Geregelt werden die Art und der Umfang der Lieferung, die Verpflichtung des dLZAs zur Archivierung, die Nutzungsbedingungen und ggf. die Kosten. Die rechtlichen Vereinbarungen werden durch konkrete Durchführungsbestimmungen ergänzt. Besteht keine Möglichkeit zu einer formalen Übereinkunft, so ist eine spezielle Begründung erforderlich.

Formale Übereinkünfte können sein:

a) Gesetze, Verordnungen: Pflichtgesetz, Archivgesetze

Gesetz über die Deutsche Nationalbibliothek, (DNBG)

<http://www.bgbportal.de/BGBL/bgb11f/bgb1106s1338.pdf>

b) Verträge, Vereinbarungen:

Lizenzverträge (vgl. Archivierungsklausel im JISC-Mustervertrag für elektronische Zeitschriften:

http://www.nesli2.ac.uk/NESLi2_licence_journals_final011003.htm

Rahmenverträge

Depositaverträge, Archivierungsvereinbarungen, Archivierungs- und Nutzungsgenehmigungen (vgl. Bayerische Staatsbibliothek: <http://www.babs-muenchen.de/content/netzpublikationen/einzelbewilligung.pdf>)

Eine entsprechende Vereinbarung bzw. die dazugehörige Durchführungsbestimmung legt u.a. fest:

i) in welcher Form die Kooperation zwischen Produzent / Lieferant und dLZA gestaltet wird; wie Rückmeldungen geregelt sind.

ii) Art und Umfang der Lieferung:

Umfang, Terminierung, Verfahren der Übergabe (Datenträger, Dateitransfer über Netze, Upload, Download), Dateiformate, weitere Eigenschaften der Dateien (z.B. ohne aktive Elemente), zusätzliche Informationen (z.B. den Inhalt und die Struktur beschreibende Metadaten, XML-Schema etc).

iii) Verpflichtung des dLZAs:

Zeitpunkt der Übernahme der rechtlichen Verantwortung, Dauer der Archivierung, Anwendung von Erhaltungsmaßnahmen (mehrfache Kopien, verändernde Eingriffe z.B. bei Migrationen), signifikante Eigenschaften.

iv) Nutzungsbedingungen:

Zielgruppen, angebotene Dienstleistungen, Nutzungsrechte, Kosten.

Bei der Archivierung der Unterlagen der STASI ist keine Möglichkeit zu einer formalen Übereinkunft gegeben, da weder der Rechteinhaber noch ein rechtlicher Nachfolger existiert.

3.2 Das digitale Langzeitarchiv handelt bei der Archivierung auf der Basis rechtlicher Regelungen.

Das dLZA berücksichtigt bei der Archivablage und dem Einsatz von Langzeiterhaltungsmaßnahmen gesetzliche Bestimmungen und vertragliche Regelungen.

Den Einschränkungen bei der Archivierung durch das Urheberrecht kann z.B. durch die explizite Vereinbarung des Rechts auf mehrfache Speicherung,

verändernde Eingriffe etc. entgegengewirkt werden.

3.3 Das digitale Langzeitarchiv handelt bei der Nutzung auf der Basis rechtlicher Regelungen.

Das dLZA berücksichtigt bei der Nutzung der digitalen Objekte gesetzliche Bestimmungen und vertragliche Regelungen. Resultieren daraus Einschränkungen bei der Nutzung, wird dokumentiert, auf welcher Grundlage diese beruhen.

Zu den rechtlichen Regelungen, die die Nutzung beeinflussen können, gehören insbesondere das Urheberrecht, der Datenschutz, andere gesetzliche Regelungen (z.B. die Schutzfristen im Bereich der Archive), vertragliche Regelungen oder die vertragliche oder gesetzliche Zweckbindung der Nutzung.

Den Einschränkungen bei der Nutzung kann zum Teil durch einen kontrollierten Zugang zu den digitalen Objekten begegnet werden. Für die Beachtung von Urheberrechten können dies die Registrierung/Anmeldung, die ausschließliche Nutzung vor Ort bzw. über das Intranet oder kostenpflichtige Nutzungs-/Abrechnungsmodelle sein. Für die Einhaltung von datenschutz- und archivrechtlichen Bestimmungen können gesonderte Verpflichtungserklärungen oder die Vorlage anonymisierter Benutzungskopien eine Möglichkeit sein.

4 Die Organisationsform ist für das digitale Langzeitarchiv angemessen.

Das dLZA ist so organisiert, dass es kurz-, mittel- und langfristig seine Ziele erfüllen kann. Seine Leistungsfähigkeit und Nachhaltigkeit sind für den Nutzer und den Produzenten bewertbar. Als Grundlage für diese Bewertung dienen die im Folgenden aufgeführten Punkte.

Erpanet-Tagung: "Business Models related to Digital Preservation", 2004. [16]

4.1 Die Finanzierung des digitalen Langzeitarchivs ist sichergestellt.

Das dLZA kann glaubhaft machen, dass die zugesicherten Leistungen momentan und entsprechend der langfristigen Planung finanziert werden können.

Es besteht eine gesetzliche Regelung, aus der sich die Grundlagen einer Finanzierung des digitalen Langzeitarchivs herleiten lassen.

Die Finanzierung ist bei staatlich finanzierten digitalen Langzeitarchiven in den formalen Planungsunterlagen (zumindest mittelfristig) enthalten.

Ein privates dLZA kann aufgrund einer kostenpflichtigen Nutzung seiner Leistungen und auf Basis eines langfristigen Geschäftsplans Aussagen zur finanziellen Nachhaltigkeit treffen.

Digital Longevity Department: Vers van de pers...Kostenmodel digitale

bewaring (Kostenmodell für die Langzeitarchivierung), 2006. [17]
Palm, Jonas: The Digital Black Hole, 2006. [18]
Oltmans, Erik; Kol, Nanda: A Comparison Between Migration and Emulation in Terms of Costs, 2005. [19]

4.2 Dem digitalen Langzeitarchiv steht Personal mit angemessener Qualifikation in ausreichendem Umfang zur Verfügung.

Die Qualifikation des Personals ist den Zielen, Aufgaben und Prozessen des dLZAs angemessen. Eine entsprechende Personalentwicklung sorgt dauerhaft für die nötige Qualifizierung und Weiterbildung. Der Umfang des Personals ist so bemessen, dass alle notwendigen Prozesse vollständig abgearbeitet werden können. Die langfristige Planung des dLZAs schließt Personalressourcen mit ein.

Eine Personalentwicklung umfasst die aufgabenorientierte Ausbildung und Weiterbildung des Personals, z.B. durch Schulung und Bereitstellung von Fachliteratur.

Als Teil der Qualifizierung dient eine aktive Teilnahme an einschlägigen nationalen und internationalen Konferenzen und Arbeitskreisen sowie die Mitarbeit in Standardisierungsgremien. Durch die aktive Teilnahme wird die Qualifikation nach außen sichtbar.

Fehlende eigene Kapazitäten können durch externe Kapazitäten ausgeglichen werden.

4.3 Für das digitale Langzeitarchiv bestehen angemessene Organisationsstrukturen.

Die Organisationsstruktur ist den Zielen, Aufgaben und Prozessen des dLZAs angemessen. Die Prozesse sowie die Zuordnung von Personal- und Sachressourcen sind so strukturiert, dass die definierten Ziele erreicht werden können.

Das dLZA ist an geeigneter Stelle im Geschäftsverteilungsplan aufgeführt. Ein Organigramm verdeutlicht diese Einordnung.

4.4 Das digitale Langzeitarchiv betreibt eine langfristige Planung.

Das dLZA betreibt eine vorsorgende Planung, in der die anstehenden oder zu erwartenden Aufgaben und die Zeitpunkte ihrer Realisierung genannt werden. Das Management verfügt über geeignete Strukturen und Verfahren für eine strategische Planung. Basis für eine langfristige Planung ist die Beobachtung der rechtlichen und gesellschaftlichen Rahmenbedingungen, der Anforderungen und Erwartungen der Zielgruppen (in OAIS: „Monitor Designated Community“) sowie aller technischen Entwicklungen (in OAIS: „Monitor Technology“), die für den dauerhaften Erhalt und die angemessene Nutzung der durch die digitalen Objekte repräsentierten Informationen

relevant sind. Die Planung umfasst auch die Sicherstellung der erforderlichen Ressourcen.

Relevante Gesetzgebungsverfahren sind bereits in frühen Phasen zu beobachten (z.B. Gesetz über Rahmenbedingungen für elektronische Signaturen).

Strategische Planung erfordert den Rückgriff auf verlässliche Ist-Daten. Eine Prozesskostenrechnung beispielsweise unterstützt eine langfristige Planung des Ressourcenbedarfs.

4.5 Das digitale Langzeitarchiv reagiert auf substantielle Veränderungen.

Substantielle Veränderungen sind dadurch charakterisiert, dass bei unterlassener Reaktion entweder die Ziele nicht mehr erfüllt werden können oder deren Erfüllung zumindest einem erhöhten Risiko ausgesetzt wird. Substantielle Veränderungen können sowohl technischer als auch organisatorischer oder gesellschaftlicher Art sein.

Das Management enthält hierzu ein Prozesselement, das Veränderungen beobachtet und mögliche Auswirkungen auf die Aufgabenerfüllung bewertet sowie erforderliche Änderungen plant, umsetzt und die Umsetzung kontrolliert.

Die Beobachtung technischer Entwicklungen umfasst beispielsweise die Entwicklung und Standardisierung neuer Dateiformate und neuer Speichertechniken sowie die dadurch ausgelöste oder zu erwartende Überalterung bestehender Techniken.

Eine substantielle Veränderung technischer Art kann der grundsätzliche Wechsel der Mensch-Maschine-Kommunikation sein.

Eine die Organisation als Ganzes betreffende substantielle Änderung kann der Wegfall der Trägerinstitution und somit der finanziellen Basis sein.

4.6 Die Fortführung der festgelegten Aufgaben ist auch über das Bestehen des digitalen Langzeitarchivs hinaus sichergestellt.

Das dLZA hat auch für einen Krisenfall vorgesorgt. Die Fortführung der Aufgaben muss in einem solchen Fall in einem anderen organisatorischen Rahmen so erfolgen, dass die festgelegten Aufgaben vollständig erbracht werden können. Sollte dies nicht möglich sein, werden die Einschränkungen dokumentiert. Das dLZA trifft Vorsorge, dass ein Übergangsprozess rechtzeitig definiert, geplant und implementiert werden kann. Eine adäquate Dokumentation ist die Grundlage für den Erfolg eines eventuellen Übergangsprozesses.

Dazu gehört die Exportierbarkeit der Gesamtheit der Archivpakete (inkl. Metadaten) in eine durch den Nachfolger interpretierbare Form, um die Interpretierbarkeit der Daten und deren Authentizität sicherzustellen. Eine außen stehende bzw. übergeordnete Instanz bürgt für die Fortführung der definierten Aufgaben.

Der Fortbestand ist durch die Vereinbarung mit einer vergleichbaren

Organisation geregelt.

5 Das digitale Langzeitarchiv führt ein angemessenes Qualitätsmanagement durch.

Das Qualitätsmanagement sorgt dafür, dass die Ziele des dLZAs erreicht werden. Dafür werden die allgemeinen Zielvorgaben in konkrete Ziele und Aufgaben herunter gebrochen. Für diese werden geeignete Prozessstrukturen festgelegt, die vom Qualitätsmanagement überwacht werden.

Das Qualitätsmanagement ist ein Querschnittsprozess, der sich auf alle Bereiche des dLZAs bezieht.

ISO 9000:2005: Quality management systems. Fundamentals and vocabulary, 2005. [20]

Liggemeyer, Peter: Software-Qualität, 2002. [21]

Kneuper, Ralf: Verbesserung von Softwareprozessen mit Capability Maturity Model Integration, 2006. [22]

ITIL: IT - Infrastructure Library. [23]

5.1 Alle Prozesse und Verantwortlichkeiten sind definiert.

Das Qualitätsmanagement sorgt dafür, dass alle Prozesse und deren Wechselwirkung definiert sind und dass insbesondere allen Prozessen Verantwortliche zugeordnet sind. Dies gilt auch für organisationsexterne (ausgelagerte) Prozesse.

Die Vollständigkeit der Prozesse und deren Wechselwirkung kann leichter festgestellt werden, wenn ein passendes Referenzmodell zur Verfügung steht. Als Ausgangsbasis zur Definition von Kernprozessen können die funktionalen Entitäten nach OAIS - Aufnahme (Ingest), Archivablage (Archival Storage) und Nutzung (Access) - herangezogen werden. Anhand dieser Kernprozesse können dann Unterstützungs- und Managementprozesse definiert werden (Datenmanagement, Qualitätsmanagement, etc.).

Organisationsexterne Prozesse erfordern z.B. als internen Prozess eine vertragsfeste Formulierung der Leistungen sowie einen Prozess, der diese Leistungen kontrolliert. Diesen internen Prozessen ist eine Verantwortlichkeit zuzuordnen.

CCSDS: Producer-Archive Interface Methodology - Abstract Standard, Blue Book, 2004. [24]

Erpanet-Tagung: "Workshop on Workflow", 2004. [25]

5.2 Das digitale Langzeitarchiv dokumentiert alle seine Elemente nach einem definierten Verfahren.

Zu den Elementen zählen: Ziele, Konzepte, Spezifikationen,

Implementationen, Prozesse, Software, Objekte und Metadaten etc.
Das Qualitätsmanagement sorgt für ein geeignetes Verfahren für die Dokumentation, d.h. für ein System zum Management aller erforderlichen Dokumente. Das dLZA legt Regeln für die Vollständigkeit, Korrektheit, Aktualität, Verständlichkeit und Verfügbarkeit der Dokumentation fest, setzt diese um und kontrolliert die Einhaltung der Regeln.
Damit wird vermieden, dass Wissen allein an Individuen gebunden ist.

Zur Verständlichkeit trägt beispielsweise eine einheitliche Sprache bei, die den Nutzern der Dokumentation angepasst ist. Die Dokumentation kann demzufolge formal (z.B. zur Beschreibung kritischer Softwareprozesse), halbformal (zur konzeptionellen Beschreibung der Prozesse und IT-Infrastruktur) oder natürlich sein (z.B. zur Außendarstellung der Ziele des Archivs).

- Softwaredokumentation
- Prozessdokumentation
- Dokumentation der Objektformate

B. Umgang mit Objekten

Das digitale Langzeitarchiv analysiert seine Ziele und Strategien und spezifiziert alle objektbezogenen Anforderungen für den Umgang mit digitalen Objekten während des Lebenszyklus der Objekte im dLZA. Dieser entspricht in den Hauptphasen dem OAIS-Referenzmodell mit den Prozessen (in OAIS: „Funktionalen Entitäten“) Aufnahme (Ingest), Archivablage (Archival Storage, inklusive Umsetzung der Langzeiterhaltungsmaßnahmen) und Nutzung (Access). Eine Erweiterung dieser Funktionen kann durch die Ziele des digitalen Langzeitarchivs erforderlich werden. Darüber hinaus orientiert sich der Umgang mit Objekten am Informationsmodell des OAIS-Referenzmodells und definiert entsprechende Übergabepakete (Submission Information Package, SIP), Archivpakete (Archival Information Package, AIP) und Nutzungspakete (Dissemination Information Package). [26]

Integrität und Authentizität der zu erhaltenden Informationen sind Kernkonzepte der Vertrauenswürdigkeit. Integrität und Authentizität sind daher nahtlos in allen Phasen zu sichern, für die das dLZA die Verantwortung übernommen hat. Voraussetzung hierfür ist die Sicherung der Integrität und Authentizität der digitalen Objekte, die die zu erhaltende Information repräsentieren (vgl. 6 und 7). Für den dauerhaften Erhalt der Informationen betreibt das dLZA eine objektbezogene Planung der Langzeitarchivierungsmaßnahmen (in OAIS: Preservation Planning) (vgl. 8). Spezifizierte Vorgaben für die Aufnahme, Archivablage und Nutzung sowie Vorgaben für die jeweiligen Objekte und ihre Transformationen sind weitere Indikatoren für Vertrauenswürdigkeit (vgl. 9, 10 und 11). Um aus den digitalen Objekten für die definierten Zielgruppen nutzbare Information zu rekonstruieren und um die Forderungen nach Integrität, Authentizität und rechtmäßiger Nutzung der Information zu erfüllen, betreibt das dLZA ein nachvollziehbares, auf Metadaten basierendes Datenmanagement (vgl. 12).

Die Forderungen aus dem Umgang mit Objekten sind Vorgabe für die Planung und den Betrieb der technischen Infrastruktur sowie des Sicherheitskonzepts (vgl. 13, 14).

6 Das digitale Langzeitarchiv stellt die Integrität der digitalen Objekte auf allen Stufen der Verarbeitung sicher.

Unter Integrität wird hier erstens die Vollständigkeit der digitalen Objekte sowie zweitens deren Unversehrtheit verstanden.

Maßstab für die Integrität sind die als erhaltenswert definierten Eigenschaften

eines digitalen Objekts (vgl. 9.2).

Risiken für die Integrität stellen einerseits menschliche (böswillige oder irrtümliche) Aktivitäten sowie andererseits technische Unvollkommenheit sowie Beschädigung oder Entwendung von technischer Infrastruktur dar.

Für die Sicherung der Integrität trifft das dLZA sowohl organisatorische als auch technische Vorkehrungen.

Das dLZA betreibt ein für den Erhalt der Integrität geeignetes Datenmanagement für die Prozesse Aufnahme (Ingest), Archivablage (Archival Storage) und Nutzung (Access). Auch für die Integrität des Datenmanagements selbst trifft das dLZA Vorkehrungen.

In Ausnahmefällen kann die Integrität zwar gestört sein, dies muss dann aber ausreichend dokumentiert werden.

Ein Beispiel für böswillige oder unbeabsichtigte Modifikation ist das Einliefern von virenbehafteten Objekten, deren Ausführung die Änderung oder Modifikation von Objekten oder weiteren Systemelementen mit sich bringt (z.B. Datenbankskripts, die Objekte oder Metadaten löschen). Beispiele für technische Unvollkommenheiten sind fehlerhafte oder unvollständige Software, insbesondere bei komplexen Transformationen (Migrationen), sowie überalterte oder nicht der Spezifikation entsprechend gelagerte Speichermedien. Im Allgemeinen werden jedoch technische Unvollkommenheiten, die absehbar sind, durch entsprechende Verfahren der Fehlerkorrektur oder Fehlererkennung bis zu einem bestimmten Grad ausgeglichen oder signalisiert. Teilweise kann der Nutzer für bestimmte Systemkomponenten höherwertigere Fehlerkorrekturverfahren (z.B. durch einen höheren Grad an Redundanz) auswählen. Diese sind ggf. auszuschöpfen.

Bei der Archivierung von Webseiten kann die Integrität der Objekte aufgrund der Funktionalitäten der verwendeten Offlinebrowser gestört sein (z.B. Fehlen bewegter Bilder), ein Protokoll dokumentiert jedoch diesen Umstand.

ISO 15489-1:2001: Information and documentation – Records Management, 2001. [27]

Shirey, R.: Internet Security Glossary, 2000. [28]

ISO/IEC 15408-x:2005: Information technology. Security techniques - Evaluation criteria for IT security, 2005. [29]

DigiCult: Integrity and Authenticity of Digital Cultural Heritage Objects, 2002. [30]

6.1 Aufnahme (Ingest): Das digitale Langzeitarchiv sichert die Integrität der digitalen Objekte.

Das dLZA spezifiziert hierzu eine klare Schnittstelle zum Produzenten und zur Archivablage. Diese schließt die Transformation von Übergabepaketen zu Archivpaketen ein. Die Schnittstelle ermöglicht den Produzenten und der Administration des dLZAs, die Integrität der digitalen Objekte zu überprüfen

und zu erhalten.

Das dLZA hat beispielsweise mit seinen Produzenten bzw. Lieferanten eine Vereinbarung getroffen, wie die Transaktionen im Zusammenhang mit der Einlieferung technisch zu realisieren sind. Insbesondere regelt eine Vereinbarung den Übergang der Verantwortlichkeit bezüglich der Integrität der Objekte.

Das dLZA hat mit dem Produzenten bzw. Lieferanten vereinbart, welche Eigenschaften digitale Objekte haben müssen, um eine Gefährdung der Integrität auszuschließen. Dies kann z.B. die vorherige Entfernung ausführbarer Codes in digitalen Dokumenten umfassen.

Das dLZA sorgt für sichere Übertragungswege vom Produzenten bzw. Lieferanten zum dLZA.

Das dLZA führt Vollständigkeits- und Qualitätskontrollen der Lieferungen durch.

CCSDS: Producer-Archive Interface Methodology - Abstract Standard, Blue Book, 2004. [24]

Littman, Justin: A Technical Approach and Distributed Model for Validation of Digital Objects, 2006. [31]

6.2 Archivablage (Archival Storage): Das digitale Langzeitarchiv sichert die Integrität der digitalen Objekte.

Das dLZA spezifiziert hierzu alle Funktionen einer Archivablage, die für die Überprüfung und den Erhalt der Integrität der digitalen Objekte durch die Administration des dLZAs notwendig sind. Die Funktionen umfassen die Abbildung der Archivpakete auf Speichermedien, die dauerhafte Speicherung, die Wiederherstellung der Archivpakete sowie alle Änderungen an den Archivpaketen.

Das dLZA hat ein nachvollziehbares Verfahren zur Bestimmung des erforderlichen Grades der physischen Redundanz sowie zur Bestimmung geeigneter Lokationen von Speichermedien sowie von zugehörigen Subsystemen.

Das dLZA legt die erforderliche Qualität der Speichermedien fest (z.B. durch die Verwendung normierter und zertifizierter Speichermedien).

Das dLZA hat eine Politik für den logischen Zugang zur Archivablage festgelegt; dies schließt dLZA-interne Nutzer wie z.B. Systemadministratoren mit ein.

Das dLZA hat den physischen Zugang zu den IT-Systemen eindeutig geregelt.

Das dLZA hat Regeln zur Auffrischung (Refreshment) der Medien bzw. zur Migration der digitalen Objekte auf andere Medien.

6.3 Nutzung (Access): Das digitale Langzeitarchiv sichert die Integrität der digitalen Objekte.

Das dLZA spezifiziert hierzu eine klare Schnittstelle zum Nutzer und zur Archivablage. Diese schließt die Transformation von Archivpaketen zu Nutzungspaketen mit ein. Die Schnittstellen ermöglichen dem Nutzer und der Administration des dLZAs, die Integrität der digitalen Objekte zu überprüfen und zu erhalten.

Das dLZA trägt Sorge dafür, dass ein unbefugter Nutzer keine Rechte über digitale Objekte, Metadaten oder sonstige Systembestandteile erhalten kann.

Das dLZA definiert, wie weit seine Verantwortlichkeit für die Integrität der digitalen Objekte beim Auslieferungsprozess reicht.

Das dLZA analysiert die Qualität seiner zur Verfügung gestellten Interpretationshilfen und macht die Ergebnisse den Nutzern bekannt.

7 Das digitale Langzeitarchiv stellt die Authentizität der digitalen Objekte auf allen Stufen der Verarbeitung sicher.

Authentizität bedeutet hier, dass das Objekt das darstellt, was es vorgibt darzustellen. Ein wichtiger Aspekt ist, dass das vorliegende Objekt von der angegebenen Quelle und zur angegebenen Zeit erstellt wurde.

Ferner schließt Authentizität den lückenlosen Nachweis aller im Sinne der Erhaltungsmaßnahmen durchgeführten Transformationen an den Objekten mit ein.

Das dLZA dokumentiert, wenn bei einem Objekt die Authentizität nicht festgestellt werden kann. Mit dem Eingang des Objektes in das dLZA übernimmt dieses die Verantwortung für die Authentizität.

Das dLZA betreibt ein für den Erhalt der Authentizität geeignetes Datenmanagement für die Prozesse Aufnahme (Ingest), Archivablage (Archival Storage) und Nutzung (Access). Dieses stellt insbesondere die Dokumentation aller Veränderungen an den Objekten (inkl. Metadaten) sicher (siehe 12.4).

Authentizität beinhaltet, dass der Erzeuger oder Sender sowie die angegebene Erstellungs- oder Sendezeit den Tatsachen entsprechen, z.B. dass eine angeblich von einer Person erstellte, zu einem gewissen Zeitpunkt versendete E-Mail tatsächlich von dieser Person stammt und zum angegebenen Zeitpunkt versendet wurde.

ISO 15489-1:2001: Information and documentation – Records Management, 2001. [27]

Shirey, R.: Internet Security Glossary, 2000. [28]

ISO/IEC 15408-x:2005: Information technology. Security techniques - Evaluation criteria for IT security, 2005. [29]

PREMIS Working Group: Data Dictionary for Preservation Metadata, Version 2.0 2008. [32]

Gladney, H. M.; Bennett, J. L. : What Do We Mean by Authentic? What's the Real McCoy?, 2003. [33]

InterPARES. [34]

7.1 Aufnahme (Ingest): Das digitale Langzeitarchiv sichert die Authentizität der digitalen Objekte.

Das dLZA spezifiziert Verfahren zur Beurteilung und Sicherung der Authentizität der Übergabepakete.

Das dLZA verlangt die formale Registrierung des Produzenten bzw. Lieferanten bei einer autorisierten Stelle.

In bestimmten Kontexten kann durch die Verwendung digitaler Signaturen die Authentizität der Übergabepakete sichergestellt werden.

Das dLZA verlangt vom Produzenten bzw. Einlieferer Verfahren, um die Authentizität der digitalen Objekte zu bewerten, z.B. anhand von Metadaten, die die Herkunft beschreiben.

CCSDS: Producer-Archive Interface Methodology - Abstract Standard, Blue Book, 2004. [24]

7.2 Archivablage (Archival Storage): Das digitale Langzeitarchiv sichert die Authentizität der digitalen Objekte.

Das dLZA spezifiziert Verfahren, die bei der Durchführung der Langzeiterhaltungsmaßnahmen die Authentizität der Objekte sicherstellen bzw. den Grad an Authentizität dokumentieren (vgl. 10.4 und 12.4).

Das dLZA führt einen lückenlosen Nachweis über alle durchgeführten manipulierenden (d.h. verändernden oder löschenden) Operationen an den digitalen Objekten.

7.3 Nutzung (Access): Das digitale Langzeitarchiv sichert die Authentizität der digitalen Objekte.

Das dLZA sichert die Authentizität der Nutzungspakete und ermöglicht dem Nutzer, den Grad ihrer Authentizität zu bestimmen. Zusätzlich authentifiziert sich hierzu das dLZA gegenüber dem Nutzer als Lieferant der Nutzungspakete.

Das dLZA stellt dem Nutzer die Metadaten, die die Herkunft und alle Veränderungen im Archivierungsprozess dokumentieren, für eine Bewertung der Authentizität zur Verfügung.

Das dLZA registriert sich bei einer autorisierten Stelle, z.B. bei der Regulierungsbehörde für Post- und Telekommunikation, und erhält dort ein digitales Signaturschlüssel-Zertifikat, welches es zum Erzeugen digitaler Signaturen benutzt.

In bestimmten Kontexten können digitale Signaturen bei der Auslieferung der Nutzungspakete zur Sicherung der Authentizität verwendet werden. Vergleiche auch die im Projekt ArchiSafe an der Physikalisch-Technischen Bundesanstalt Braunschweig eingesetzte Methode zur Nutzung digitaler Signaturen (<http://www.archisafe.de/s/archisafe/index>).

8 Das digitale Langzeitarchiv betreibt eine langfristige Planung seiner technischen Langzeiterhaltungsmaßnahmen.

Um seiner Verantwortung für den Erhalt der Informationen gerecht zu werden, nimmt das dLZA eine langfristige Planung auf, in der die anstehenden oder zu erwartenden Aufgaben und die Zeitpunkte ihrer Realisierung genannt werden. Diese strategische Planung (vgl. 4.4) wird auf Objektebene spezifiziert. Diese Maßnahmen folgen dem kontinuierlichen Wandel der Technik (Wechsel der verfügbaren Datenträger, Datenformate, Anforderungen der Nutzer etc.).

Zum Langzeiterhalt der durch digitale Objekte repräsentierten Informationen werden Maßnahmen zum physischen Erhalt der Daten (Integrität, Authentizität) und deren Verfügbarkeit sowie zum Erhalt der Interpretierbarkeit angewendet.

Langzeiterhaltungsmaßnahmen beziehen sich sowohl auf die Inhaltsdaten als auch auf die Metadaten.

Zur Umsetzung der Langzeiterhaltungsmaßnahmen siehe 10.4.

Als Langzeiterhaltungsmaßnahme kann sich bei geeigneten digitalen Objekten die Ausgabe auf analoge Medien (z.B. Mikrofilm) und eine Redigitalisierung bei Bedarf anbieten.

Zum Erhalt der Interpretierbarkeit werden hauptsächlich folgende Methoden eingesetzt:

Konvertierung in ein aktuelles Format bzw. eine aktuelle Formatversion (Migration, z.B. Transformation von Dateien im Format PDF 1.4 in das Format PDF/A)

Nachbildung der alten Anwendungsumgebung in einer neuen technischen Infrastruktur (Emulation, z.B. DIOSCURI, <http://dioscURI.sourceforge.net/>)

Eine langfristige Planung der Aufgaben, die sich aus den Formaten ergeben, kann sich z.B. an ein Format Register anlehnen. Format Register werden derzeit u.a. von Harvard <http://hul.harvard.edu/gdfr/> und den National Archives, Kew (PRONOM: <http://www.nationalarchives.gov.uk/pronom/>) entwickelt.

Der Entscheidungsprozess kann unterstützt werden durch die Methode (und in Kürze das entsprechende Tool PLATO), das im Rahmen des Projekts PLANETS entwickelt wurde.

nestor - Handbuch : Digitale Erhaltungsstrategien, 2008. [35]

Rauch, Carl; Rauber, Andreas: Anwendung der Nutzwertanalyse zur Bewertung von Strategien zur langfristigen Erhaltung digitale Objekte, 2006. [36]

Strodl, Stephan; Becker, Christoph et al.: How to Choose a Digital Preservation Strategy: Evaluating a Preservation Planning Procedure, 2007. [37]

Oltmans, Erik; Kol, Nanda: A Comparison Between Migration and Emulation in Terms of Costs, 2005. [19]

9 Das digitale Langzeitarchiv übernimmt digitale Objekte von den Produzenten nach definierten Vorgaben.

Die allgemeinen Sammelrichtlinien, Selektionskriterien, Bewertungskriterien oder Kriterien für die Überlieferungsbildung (vgl. 1.1) sowie die allgemeinen Ziele des dauerhaften Erhalts (vgl. 1.2) werden auf Objektebene spezifiziert. Die Übernahme kann durch das Einstellen der Objekte in das dLZA durch die abgebende Stelle erfolgen oder durch manuelles bzw. automatisches Sammeln des dLZAs.

DOMEA : Aussonderung und Archivierung elektronischer Akten. Erweiterungsmodul zum Organisationskonzept 2.1, 2005. [38]

The U.S. National Archives & Records Administration: Disposition of Federal Records. Subpart L -- Transfer of Records to the National Archives of the United States, Part 1228, § 1228.270, 2002. [39]

NDAD: Transfer Procedures (Overview), 2005. [40]

DPC: Decision Tree for Selection of Digital Materials for Long-term Retention, 2006. [41]

nestor: Wege ins Archiv: Ein Leitfaden für die Informationsübernahme in das digitale Langzeitarchiv. [42]

9.1 Das digitale Langzeitarchiv spezifiziert seine Übergabepakete (Submission Information Packages, SIPs).

Das dLZA vereinbart mit den Produzenten bzw. Lieferanten bzw. legt fest, welche digitalen Objekte und Metadaten (in einem Übergabepakete als konzeptionelle Einheit) ins dLZA aufgenommen werden. Diese Vereinbarungen ermöglichen es, die Übergabe bzw. das Sammeln zu automatisieren sowie Workflows für die Aufnahme ins dLZA zu implementieren.

Die Spezifikation ist Grundlage der Qualitätsprüfung der Übergabeobjekte.

Übergabepakete können neben Inhaltsdaten auch Metadaten, z.B. zur Feststellung ihrer Authentizität, enthalten.

Beim Harvesting mittels Offline-Browser werden nur Textdokumente mit Images, aber keine Audio-, Video- und weitere multimediale Inhalte gesammelt (durch Auswahl bzw. Ausschluss spezieller Dateiformate.)

Zur Qualitätskontrolle kann die Validierung der Dateiformate der Übergabepakete mittels JHOVE (vgl. <http://hul.harvard.edu/jhove/>) erfolgen.

Das dLZA empfiehlt Dateiformate für die Übergabepakete, z.B. GeoTIFF für Daten der Fernaufklärung oder Seed/MiniSeed als Format für Geodaten, wie in GeoFon (<http://www.gfz-potsdam.de/geofon/>) verwendet.

Im Projekt kopal werden die Übergabepakete im Universellen Objektformat (UOF) spezifiziert.

Beispiele und die Spezifikation finden sich bei:

http://kopal.langzeitarchivierung.de/index_objektspezifikation.php.de

9.2 Das digitale Langzeitarchiv identifiziert, welche Eigenschaften der digitalen

Objekte für den Erhalt von Information signifikant sind.

Bei der Entscheidung über den Umfang der zu bewahrenden Eigenschaften ist vor dem Hintergrund der eigenen Ziele zwischen den technischen Möglichkeiten sowie dem Aufwand für die Langzeitarchivierung einerseits und den Bedürfnissen der Zielgruppe(n) andererseits abzuwägen.

Es kann effektiv sein, verschiedene Repräsentationen eines Informationsobjektes zu erhalten, um möglichst viele Eigenschaften zu bewahren.

Bei Informationen aus Datenbanken kann es ausreichen, nur die Daten als so genannte „flat files“ zu archivieren (einschließlich einer genauen Beschreibung der Datenstruktur).

Bei elektronischen Akten werden nach den Vorgaben des DOMEA-Konzepts die einzelnen Dokumente als Bilddateien gespeichert. Dabei geht die Möglichkeit der Volltextrecherche sowie die Ausführbarkeit einiger Dokumente (Excel-Tabellen oder Powerpoint-Präsentationen) verloren.

Von Internetseiten mit Text-Bild-Informationen kann das eine Archiv ausschließlich die Textinformationen, das zweite nur die Bilder und das dritte den gesamten Zusammenhang archivieren. Die unterschiedlichen Zielsetzungen münden in entsprechend unterschiedlichen Archivierungsstrategien.

Von Internetseiten werden einerseits Screenshots eines gängigen Browsers aufbewahrt, andererseits die Textinformationen für eine komfortable Suche.

Kunze, John: Future-Proofing The Web: What We Can Do Today, 2005. [43]

9.3 Das digitale Langzeitarchiv erhält die technische Kontrolle über die digitalen Objekte, um Langzeitarchivierungsmaßnahmen durchführen zu können.

Viele digitale Objekte enthalten technische Vorkehrungen, die eine Nutzung aus kommerziellen oder rechtlichen Gründen einschränken. Für die langfristige Archivierung digitaler Objekte ist es unabdingbar, dass das dLZA die Objekte uneingeschränkt öffnen und bearbeiten kann. Sämtliche technische Einschränkungen der Benutzbarkeit müssen daher vor der Abgabe ans dLZA entfernt werden.

Interne Einstellungen verhindern z.B. das Kopieren, Drucken oder Speichern der Objekte; andere Objekte sind verschlüsselt und verlangen die Eingabe von Kennwörtern oder sind nach Ablauf eines bestimmten Zeitraums oder nach einer definierten Zahl an Nutzungen nicht mehr zu öffnen.

„Musikwirtschaft und Buchbranche vereinbaren mit der deutschen Nationalbibliothek Vervielfältigung kopiergeschützter Werke“, Gemeinsame Presseinformation der Deutschen Bibliothek, des Börsenvereins des Deutschen Buchhandels und des Bundesverbands der Phonographischen Wirtschaft

http://www.ddb.de/aktuell/presse/pressemitte_vervielfaeltigung.htm

10 Die Archivablage der digitalen Objekte erfolgt nach definierten Vorgaben.

Der Kern eines digitalen Langzeitarchivs liegt in der Durchführung des eigentlichen Archivierungsprozesses. Dieser umfasst die Definition der Archivpakete, die Speicherung der Archivpakete sowie die Umsetzung der Langzeiterhaltungsmaßnahmen.

10.1 Das digitale Langzeitarchiv definiert seine Archivpakete (Archival Information Packages, AIPs).

Archivpakete sind konzeptionelle Einheiten, die aus den Inhaltsdaten und allen für die Langzeitarchivierung relevanten Metadaten (vgl. 12) bestehen. Die Definition der Archivpakete schließt die Festlegung der genutzten Paket- und Objektstrukturen, sowie geeigneter Speicherorte und -formate mit ein. Die Wahl der Archivpakete hängt von den Objekttypen (z.B. digitales Textbuch oder 3D-Animation) und den zu erhaltenden Eigenschaften der Objekte ab.

Für die Archivdateiformate sind offene bzw. offen gelegte Formate und häufig verwendete Formate zu bevorzugen, da hier erstens von einer längeren Lebensdauer ausgegangen werden kann und zweitens das Vorhandensein von Verfahren und Werkzeugen zur Konvertierung bzw. Emulation wahrscheinlicher ist, da diese Formate durch einen breiten Anwenderkreis unterstützt werden.

Beispiele für aktuell verwendete Archivformate, deren Angemessenheit im Einzelfall zu überprüfen ist:

- für unformatierten Text: ASCII/Unicode
- für strukturierte Texte: XML (<http://www.w3.org/XML/>)
- für formatierte Texte: PDF/A (ISO 19005-1: 2005, http://www.iso.org/iso/catalogue_detail?csnumber=38920)
- für Rastergrafiken: TIFF 6.0 (<http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>)
- für Audioformate: WAVE (http://msdn.microsoft.com/library/default.asp?url=/library/en-us/multimed/htm/_win32_waveformat_str.asp)
- für Videodateien: MPEG 4 File Format (ISO/IEC 14496, <http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm>)
- für ausführbare Programme: deren Quelltext und eine Dokumentation der Programmiersprache

Bei der Entscheidung für eine (verlustfreie) Komprimierung der Daten ist zwischen dem Speicherplatzgewinn und einer weiteren Abhängigkeit vom Komprimierungsverfahren abzuwägen. Offene bzw. offen gelegte Verfahren sind vorzuziehen.

Zur Strukturbeschreibung der Archivpakete bietet sich zur Zeit XML an, insbesondere nach dem METS-Schema, das es erlaubt, die Metadaten und Referenzen auf die einzelnen Dateien eines Objekts oder die Dateien selbst zu verwalten.

- Coy, Wolfgang: nestor - materialien 5: Perspektiven der Langzeitarchivierung multimedialer Objekte 2006. [44]*
- Witthaut, Dirk; Zierer, Andrea et al.: nestor - materialien 2: Digitalisierung und Erhalt von Digitalisaten in deutschen Museen, 2005. [45]*
- Erpanet-Tagung: "File Formats for Preservation", 2004. [46]*
- Abrams, Stephen: Digital Formats And Preservation, 2005. [47]*
- LOC: Sustainability of Digital Formats. Planning for Library of Congress Collections, 2006. [48]*
- ISO 19005-1:2005: Document management - Electronic document file format for long-term preservation, 2005. [49]*
- Helfer, Bernward; Lupprian, Karl-Ernst: Dateiformate: Eigenschaften und Eignung für die Archivierung elektronischer Unterlagen. Eine Handreichung für Archivarinnen und Archivare, 2004. [50]*
- BSI: IT-Grundschutz-Kataloge : Auswahl geeigneter Datenformate für die Archivierung von Dokumenten, 2007. [51]*
- Gutzmann, Ulrike; Kamp, Ulrich et al.: Praktische Lösungsansätze zur Archivierung digitaler Unterlagen: "Langzeitarchivierung" und dauerhafte Sicherung der digitalen Überlieferung, 2007. [52]*
- AK Elektronische Archivierung: Matrix zur Bewertung von Dateiformaten, 2006. [53]*
- Van Wijk, Caroline; Rog, Judith: Evaluating File Formats for Long-term Preservation, 2007. [54]*
- Steinke, Tobias: Universelles Objektformat: Ein Archiv- und Austauschformat für digitale Objekte, 2006. [55]*

10.2 Das digitale Langzeitarchiv sorgt für eine Transformation der Übergabepakete in Archivpakete.

Beim Prozess der Aufnahme ins dLZA werden die Übergabepakete in Archivpakete überführt und spezifische, für die Langzeitarchivierung notwendige Metadaten hinzugefügt. Dabei wird ggf. eine Formatkonversion durchgeführt.

Doc-Dateien werden nach PDF/A konvertiert.

10.3 Das digitale Langzeitarchiv gewährleistet die Speicherung und Lesbarkeit der Archivpakete.

Das dLZA stellt durch entsprechende Methoden sicher, dass die Archivpakete mit im System zur Verfügung stehenden Mitteln korrekt gespeichert und gelesen werden können. Unter Lesbarkeit wird die Lesbarkeit der Speichermedien und der Bitfolge verstanden.

Zur Sicherung der Integrität der Archivpakete siehe 6.2.

Möglichkeiten zur Speicherung und Sicherung der Lesbarkeit sind:

- Einsatz von RAID-Systemen
- digitale Speicherung auf geeigneten Medien wie Bänder, Platten, CDs, DVDs
- analoge Speicherung auf Mikrofilm. Die Benutzer greifen entweder auf analoge Bilder auf dem Film oder auf digitale Daten nach vorhergehender Redigitalisierung zu (Konversionsstrategie)

Vgl. dazu auch die im ARCHE-Projekt (http://www.landesarchiv-bw.de/sixcms/detail.php?template=hp_artikel&id=4802&id2=6625&sprache=de) entwickelten Methoden der Langzeitspeicherung.

10.4 Das digitale Langzeitarchiv setzt Strategien zum Langzeiterhalt der Archivpakete um.

Die gemäß Punkt 8 spezifizierten Langzeiterhaltungsmaßnahmen werden umgesetzt. Für jedes Archivpaket ist ein Zeitpunkt oder ein Anlass definiert, zu dem geprüft wird, ob eine Langzeiterhaltungsmaßnahme – z.B. ein Migrationsschritt oder die Bereitstellung einer Emulationssoftware – erfolgen muss. Bei Bedarf wird eine entsprechende Maßnahme durchgeführt und nachgewiesen (vgl. 12.4).

Eine derartige Strategie wäre zum Beispiel die Festlegung, dass 2007 überprüft wird, ob die gespeicherten Dokumente in PDF1.1 nach PDF/A migriert werden müssen.

11 Das digitale Langzeitarchiv ermöglicht die Nutzung der digitalen Objekte nach definierten Vorgaben.

Die in Punkt 2 beschriebenen Ziele der Nutzung müssen auf Objektebene spezifiziert werden. Die Nutzung kann durch Personen, aber auch durch Client-Systeme erfolgen. Die Möglichkeiten der Recherche und des Zugriffs auf die Nutzungspakete werden definiert. Jede Recherche mündet in eine eindeutige Antwort des Systems. Wenn das dLZA Teil eines größeren Archivs ist, müssen zusammengehörende digitale und analoge Objekte auch in ihrem Zusammenhang dargestellt werden. In besonderem Maße gilt dies für die Bestandteile von Hybridobjekten.

Nutzungspakete sind diejenigen Informationseinheiten, welche die Benutzer als Antwort auf ihre Anfragen ans dLZA erhalten.

11.1 Das digitale Langzeitarchiv definiert seine Nutzungspakete (Dissemination Information Packages, DIPs).

Abhängig von der(n) Zielgruppe(n) und den Archivpaketen definiert das dLZA seine Nutzungspakete. Voraussetzung dafür ist auch die Festlegung der Referenzanwendungsumgebung, in der die Objekte genutzt werden können.

Es ist möglich, dass ein Archivpaket je nach Nutzungskontext in verschiedenen Nutzungspaketen angeboten wird. Die Nutzung der durch die digitalen Objekte repräsentierten Information bedeutet in den meisten Fällen nicht den Zugriff auf die Archivpakete selbst, sondern die Nutzung von Kopien oder Derivaten (ggf. zusammen mit weiteren Informationen), die der Interpretierbarkeit dienen. Dies kann eine technische Beschreibung, zusätzliche Anwendungssoftware oder Emulationssoftware sein.

Zum Datenaustausch mit anderen digitalen Langzeitarchiven oder bei der Migration zu einer anderen technischen Infrastruktur ist es notwendig, Teile oder den gesamten Inhalt des dLZAs in ein dokumentiertes, standardisiertes Exportformat zu transformieren. Damit kann der Erhalt der Information über die Lebensdauer des dLZAs hinaus gesichert werden (vgl. 4.5).

Bildarchiv: für die Nutzung im Web werden aus den Master-Images Dateien in niedriger Auflösung erzeugt, die von heutigen Browsern dargestellt werden können. Für Reproduktionszwecke können hochauflösende Dateien elektronisch ausgeliefert werden.

Vgl. <http://www.bsb-muenchen.de/karten/bilddatenb.htm>

Im Projekt kopal werden die Nutzungspakete im Universellen Objektformat (UOF) spezifiziert. Beispiele und die Spezifikation finden sich bei: http://kopal.langzeitarchivierung.de/index_objektspezifikation.php

11.2 Das digitale Langzeitarchiv gewährleistet eine Transformation der Archivpakete in Nutzungspakete.

Die Nutzungspakete werden nach einem definierten Verfahren aus den Archivpaketen abgeleitet. Nutzungspakete können im digitalen Langzeitarchiv vorgehalten und bei veränderten Bedingungen neu produziert werden oder nur im Nutzungsfall direkt aus dem Archivpaket erzeugt werden (on the fly).

Bei der Konvertierung hochauflösender Master-Images zu Nutzungsvarianten in niedriger Auflösung, die von gängigen Browsern dargestellt werden können, sollten Informationen über den Konvertierungsvorgang festgehalten werden (Konvertierungssoftware, Parameter).

12 Das Datenmanagement ist dazu geeignet, die notwendigen Funktionalitäten des digitalen Langzeitarchivs zu gewährleisten.

Das Datenmanagement ist ein Querschnittsprozess, der die Kernprozesse eines dLZAs – Aufnahme, Archivablage und Nutzung – sowie die Planung und Durchführung der Langzeiterhaltungsmaßnahmen unterstützt und dabei die Integrität und Authentizität auf allen Stufen der Verarbeitung sicherstellt. Der Umfang des Datenmanagements orientiert sich an den Zielen des dLZAs. Folgende Aufgaben muss das Datenmanagement erfüllen:

- zur Verwaltung der digitalen Objekte und ihrer Beziehungen ist die Identifizierung der Objekte notwendig,
- die Auffindbarkeit und Nutzung setzt eine formale, inhaltliche und strukturelle Beschreibung der digitalen Objekte voraus,
- die Gewährleistung der Interpretierbarkeit, die Sicherung der Integrität sowie die Planung und Durchführung von Langzeiterhaltungsmaßnahmen setzt eine technische Beschreibung der Objekte voraus,
- zur Sicherung der Authentizität der Daten ist eine Dokumentation aller Veränderungen an den digitalen Objekten notwendig,
- zur Einhaltung der rechtlichen Vorgaben während der gesamten Verarbeitung ist die Verzeichnung aller rechtlichen Einschränkungen sowie deren Grundlage (Gesetze, Verordnungen, Verträge, Vereinbarungen) notwendig.

Diese Aufgaben werden zurzeit durch die Erhebung und Speicherung von Metadaten erfüllt. Metadaten können in einem Metadatenschema strukturiert erfasst werden. Es haben sich verschiedene Metadatenschemata für unterschiedliche Verwendungszwecke (z.B. beschreibende, strukturelle, technische, administrative, rechtliche Metadaten) und für unterschiedliche Bereiche (z.B. Archive, Bibliotheken, Museen) etabliert. Die Orientierung an einem nationalen oder internationalen Standard bzw. die Nachnutzung eines weit verbreiteten Metadatenschemas ist oft möglich und im Hinblick vor allem auf die Nachhaltigkeit, aber auch auf Kooperationen und den Datenaustausch zwischen Produzenten bzw. Lieferanten, dLZA und Nutzern sinnvoll. Ein Metadatenschema enthält definierte Felder (Datenelemente), in denen die jeweiligen Inhalte erfasst werden. Dadurch entsteht eine gleichermaßen für Menschen wie Maschinen verwertbare Datenstruktur.

Das dLZA legt Regeln für die inhaltliche Füllung der Felder fest (z.B. Nutzung von kontrolliertem Vokabular). Verschiedene Werkzeuge ermöglichen die automatische Generierung bzw. Extraktion von Metadaten (z.B. JHOVE für technische Metadaten).

In diesem Kriterienkatalog werden Metadaten als Teile der konzeptionellen Einheiten Übergabepaket, Archivpaket und Nutzungspaket aufgefasst. Verwaltet werden können diese z.B. in Datenbanken und/oder XML-Strukturen.

*Bischoff, Frank M.: Metadata in preservation: selected papers from an ERPANET seminar at the Archives School Marburg, 2004. [56]
METS: METS Schema 1.7 Documentation, 2007. [57]*

PREMIS Working Group: Data Dictionary for Preservation Metadata, Version 2.0 2008. [32].

LMER, 2005. [58]

nestor - Handbuch : Metadatenstandards im Bereich der digitalen Langzeitarchivierung, nationale und internationale Entwicklungen, 2008. [59]

NATLIB: Metadata Standards Framework – Preservation Metadata (Revised), 2003. [60]

12.1 Das digitale Langzeitarchiv identifiziert seine Objekte und deren Beziehungen eindeutig und dauerhaft.

Ein dLZA verwendet intern Identifikatoren zur Verwaltung der Objekte und ihrer Teile und deren Beziehungen (Teile/Gesamtheiten, verschiedene Varianten, Versionen etc.), insbesondere zur eindeutigen Zuordnung der Inhaltsdaten zu den Metadaten (vgl. 12.7).

Der Einsatz von nach außen sichtbaren, standardisierten Persistenten Identifikatoren stellt die zuverlässige Referenzierung der Objekte und damit die Zitierbarkeit sicher.

Durch Verwendung eines Resolving-Dienstes können Persistente Identifikatoren in die aktuelle URL-Adresse aufgelöst werden und damit den dauerhaften Zugriff sicherstellen. Dies erfordert laufende Datenpflege des Resolving-Dienstes.

Das dLZA verpflichtet sich, den Resolving-Dienst bei der Datenpflege zu unterstützen.

Aus der Welt der gedruckten Materialien werden auf elektronische Medien übertragen:

- Signaturen
- ISBN (International Standard Book Number) für Monografien,
- ISSN (International Standard Serial Number) für Periodika
 - ISBN (<http://www.ietf.org/rfc/rfc3187.txt>) und ISSN (<http://www.ietf.org/rfc/rfc3044.txt>) sind als URN-Namespaces registriert.

Für elektronische Medien werden weitere Systeme angewendet wie z.B.:

- Uniform Resource Names (URN, <http://www.ietf.org/rfc/rfc1737.txt>): Ein internationaler Internet-Standard zur eindeutigen, dauerhaften Identifizierung von Objekten. Im Bereich der Bibliotheken wird z.B. National Bibliography Number (NBN), ein Unternehmensraum der URNs, verwendet, z.B.
 - URN: urn:nbn:de:0008-20050117016
 - URL: <http://nbn-resolving.de/urn:nbn:de:0008-20050117016>
- das Handle-System (HDL, <http://www.handle.net/>): mit dem Handle-System werden Funktionen, welche die Vergabe, Administration und Auflösung von PIs in Form von Handles erlauben, bereitgestellt, z.B.
 - HDL: 1721.1/30592
 - URL: <http://hdl.handle.net/1721.1/30592>
- Digital Object Identifier (DOI, <http://www.doi.org/>) DOI wird im Bereich der Verlage, aber zunehmend auch für Fach- und Primärdaten angewendet. Die technische Grundlage bildet das

- Handle-System, z.B.
 - o DOI: 10.1045/april2004-dobratz
 - URL: <http://dx.doi.org/10.1045/april2004-dobratz>
- Archival Resource Key (ARK, <http://www.ietf.org/internet-drafts/draft-kunze-ark-10.txt>) Der Identifier kann unterschiedlich verwendet werden: Als Link von einem Objekt zur zuständigen Institution, von einem Objekt zu Metadaten und zu einem Objekt oder dessen adäquater Kopie, z.B.
 - o <http://foobar.zaf.org/ark:/12025/654xz321/s3/f8.05v.tiff>
- Scientific reference linking system (SRef, <http://www.sref.org/site/index.php>)

DNB: Persistent Identifier: Der Service der Deutschen Nationalbibliothek. [61]

PADI: Persistent Identifiers. [62]

Erpanet-Tagung: "Persistent Identifier", 2004. [63]

nestor - Handbuch : Persistent Identifier, 2008. [64]

12.2 Das digitale Langzeitarchiv erhebt in ausreichendem Maße Metadaten für eine formale und inhaltliche Beschreibung und Identifizierung der digitalen Objekte.

Umfang, Struktur und Inhalt der beschreibenden Metadaten sind abhängig von den Zielen des dLZAs, von der Zielgruppe des dLZAs und den Objekttypen. Eine formale und inhaltliche Beschreibung der Objekte in Form von Metadaten stellt die Auffindbarkeit sicher und ist ausschlaggebend für die Möglichkeiten, die den Nutzern bei der Recherche geboten werden.

In den verschiedenen Bereichen haben sich inzwischen mehrere Schemata etabliert:

Im Bereich der Bibliotheken:

Dublin Core <http://dublincore.org/>,

Maschinelles Austauschformat für Bibliotheken (MAB, <http://www.d-nb.de/standardisierung/formate/mab.htm>),

Machine-Readable Cataloging (MARC, <http://www.loc.gov/marc/>),

Metadata Objects Description Schema (MODS, <http://www.loc.gov/standards/mods/>).

In Verbindung damit können bibliothekarische Regelwerke zur Anwendung kommen, z.B. RAK oder AACR2 für die Formalerschließung, sowie RSWK oder eine Klassifikation (z.B. DDC, RVK) für die Sacherschließung.

Im Bereich der Archive:

General International Standard Archival Description (ISAD(G) [http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf)),

Encoded Archival Description (EAD, <http://www.loc.gov/ead/>), ergänzt durch Encoded Archival Context (EAC, <http://jefferson.village.virginia.edu/eac/>).

Für Daten mit Raumbezug: ISO Standard 19115.

NASA DIF (Data Interchange Format, <http://gcmd.nasa.gov/User/difguide/difman.html>) als Beschreibungsformat der NASA, welches sich zu einem de-facto Standard entwickelt hat, der auch für das Global Change Master Directory Anwendung findet (<http://gcmd.nasa.gov/>).

Shepherd, Elizabeth; Smith, Charlotte: The Application of ISAD(G) to the Description of Archival Datasets, 2000. [65]

DOMEA : Organisationskonzept, 2005. [66]

GDA: Metadaten für die Aussonderung und Archivierung digitaler Sachakten, 2004. [67]

12.3 **Das digitale Langzeitarchiv erhebt in ausreichendem Maße Metadaten zur strukturellen Beschreibung der digitalen Objekte.**

Die Struktur komplexer Objekte muss adäquat beschrieben werden, so dass die Objekte als Ganzes rekonstruiert und genutzt werden können.

Zur Darstellung der Strukturen von digitalen Objekten bieten sich verschiedene - häufig als XML-Schemata ausgeprägte - Standards an:

METS, TEI (Text Encoding Initiative, <http://www.tei-c.org/>), EAD.

Aber auch in den beschreibenden Metadaten sowie in den Metadaten für die Langzeitarchivierung (z.B. PREMIS und LMER) können Strukturinformationen verwaltet werden.

Ein digitaler Akt besteht in der Regel aus Vorgängen, diese wiederum aus Dokumenten, zu welchen weitere Dokumente (Anlagen) gehören können. Diese Hierarchie wird durch eine Datei beschrieben, die zu jeder Ebene Metadaten und auf der Dokumentebene Metadaten und Verweise auf die Dokumente selbst (Primärinformationen) enthält.

Die digitalisierte Version eines konventionellen Buches besteht aus 200 einzelnen Bilddateien. In den Metadaten sollte die korrekte Reihenfolge der Buchseiten bzw. der entsprechenden Bilddateien verzeichnet sein.

Eine archivierte Website besteht aus mehreren HTML-Seiten und JPEG-Bilddateien, die untereinander durch Links verknüpft sind. Diese Verknüpfungen sollten in den Metadaten entsprechend verzeichnet sein.

12.4 **Das digitale Langzeitarchiv erhebt in ausreichendem Maße Metadaten, die alle vom digitalen Langzeitarchiv vorgenommenen Veränderungen an den digitalen Objekten verzeichnen.**

Das dLZA dokumentiert alle Veränderungen an den digitalen Objekten. Dies beinhaltet auch den Nachweis der Beteiligten (Personen oder Systeme) und der entsprechenden Rechte (vgl. 3.2). Diese Maßnahme dient sowohl dem Nachweis von Authentizität (vgl. 7) als auch der technischen Erhaltung der digitalen Objekte.

Vor allem bei dLZAs, welche Migration als Strategie zur Langzeitarchivierung gewählt haben, werden die digitalen Objekte regelmäßig verändert. Hinzu kommen ggf. die Transformationen, die im Rahmen der Aufnahme ins dLZA sowie für die Auslieferung als Nutzungsobjekt durchgeführt werden.

Diese Metadaten (History, Audit Trail, Provenance) können verwaltet werden: z.B. im Rahmen von METS (Abschnitt amdSec digiprovMD), PREMIS (Abschnitt Events), LMER (Abschnitt Prozesse).

Ein Archiv migriert Objekte in einem veralteten Datenformat mittels eines Konvertierungsprogramms in ein aktuelles Format. Es werden Metadaten über den Migrationsvorgang, das technische Protokoll, den Zeitpunkt der Migration, die Beteiligten (Mitarbeiter und technische Hilfsmittel) sowie den Ausgang des Eingriffs erhoben und gespeichert.

12.5 Das digitale Langzeitarchiv erhebt in ausreichendem Maße Metadaten zur technischen Beschreibung der digitalen Objekte.

Zur Gewährleistung der Interpretierbarkeit, zur Sicherung der Integrität und zur Steuerung der Langzeiterhaltungsmaßnahmen müssen die Objekte selbst sowie bei komplexen Objekten all ihre Dateien technisch umfassend beschrieben werden.

Dazu gehört insbesondere die Beschreibung der Repräsentationsinformation.

Die technische Beschreibung beinhaltet allgemeine Informationen, die für alle Dateiformate angewendet werden können, u.a.:

- Dateinamen, Speicherort
- Dateigröße, verschiedene Prüfsummen
- die ausführliche Beschreibung der Dateiformate
- die bei der Erstellung verwendete Hardware/Software-Umgebung
- die zur Nutzung nötige Hardware/Software-Umgebung
- die Verzeichnung aller nötigen Zusatzobjekte (DTD, Schema-Datei, Schriften etc.)
- die beim Transfer in das Archiv verwendete Software (z.B. Offline-Browser)

Ferner gibt es für die einzelnen Formate spezifische Informationen, z.B. für TIFF-Dateien die Auflösung, den Farbraum, die Komprimierung etc.

Die allgemeinen technischen Metadaten werden ebenfalls im Rahmen von METS (Abschnitt amdSec, techMD) bzw. PREMIS oder LMER verwaltet.

Für formatspezifische Metadaten haben sich weitere Standards etabliert:

- für Bilder Metadata for Images in XML-Schema (MIX, <http://www.loc.gov/standards/mix/>) aufbauend auf NISO Technical Metadata for Digital Still Images,
- für Texte als Erweiterung des METS Schemas: textmd.xsd

Zur Beschreibung der Dateiformate können Datei-Format-Register referenziert werden z.B.:

Global Digital Format Registry: <http://hul.harvard.edu/gdfr/>,

PRONOM: <http://www.nationalarchives.gov.uk/pronom/>.

Zur Identifizierung von Dateiformaten stehen Werkzeuge zur Verfügung z.B. Digital Record Object Identification (DROID, <http://droid.sourceforge.net/wiki/index.php/Introduction>) darüber hinaus zur automatischen Extraktion technischer Metadaten, z.B. JSTOR/Harvard Object Validation Environment (JHOVE, <http://hul.harvard.edu/jhove/>).

Technische Metadaten müssen nicht notwendig vollständig explizit erhoben und verwaltet werden, sie können zum Teil auch erst bei Bedarf extrahiert werden. (z.B. vor einem Migrationsschritt).

Beispiele:

Ein dLZA speichert Dateien im Format PDF der Version 1.4. Zur Betrachtung der Dateien benötigt man das Programm „Acrobat Reader“ ab der Version 5.0. Dieses Programm läuft auf einem Microsoft Betriebssystem ab der Version Windows 98 SE. Die gesamte Software wiederum braucht einen Rechner mit einem Prozessor von mindestens 350 MHz und einem Arbeitsspeicher von 64 MB. Diese technischen Angaben sind Teil der Metadaten, die vom dLZA erhoben und gespeichert werden.

Ein dLZA speichert Dateien im Format PDF/A-1. Dieses Format wird in der ISO-Norm 19005-1:2005 ausführlich beschrieben. Das dLZA fügt die entsprechende ISO-Norm den Metadaten hinzu oder referenziert darauf über einen Verweis innerhalb der Metadaten.

Ein dLZA speichert Dateien im XML-Format. Um die Validität dieser Dateien beurteilen zu können, werden die entsprechenden Schema-Dateien benötigt. Das dLZA fügt die entsprechenden Schema-Dateien den Metadaten hinzu oder referenziert darauf über einen Verweis innerhalb der Metadaten.

12.6 Das digitale Langzeitarchiv erhebt in ausreichendem Maße Metadaten, die die entsprechenden Nutzungsrechte und -bedingungen verzeichnen.

Die Nutzung der digitalen Objekte kann aus gesetzlichen oder vertraglichen Gründen eingeschränkt sein. Abhängig von diesen Rahmenbedingungen und den entsprechenden Nutzergruppen müssen diese Rechte und Bedingungen so erfasst werden, dass damit einerseits die Nutzung gesteuert werden kann (z.B. kontrollierter Zugang, anonymisierte Benutzungskopie) und sie andererseits dem Nutzer bekannt gegeben werden können (vgl. 3.3).

Ein dLZA archiviert Datenbanken, die erst nach einer Sperrfrist von 60 Jahren zur Nutzung freigegeben sind. Ausgenommen ist die Nutzung für wissenschaftliche Forschungsvorhaben. Diese Nutzungseinschränkung ist Teil der Metadaten des dLZAs, zudem wird auf die relevanten Paragraphen (in diesem Fall § 2 Abs. 4 Satz 2 und § 5 Abs. 3 BArchG sowie § 16 Abs. 6-9 BStatG) verwiesen.

Dies kann z.B. im RightsDeclarationMD Extension Schema zu METS erfolgen. Autorenbasierte Rechtevergabe mit Markup, z.B. Creative Commons (<http://www.creativecommons.org>) oder wie bei DPPL (Digital Peer Publishing License), siehe <http://www.dipp.nrw.de/lizenzen>

12.7 Der Erhalt der Paketstruktur ist zu jeder Zeit gegeben.

Die Verknüpfung der Metadaten mit den Inhaltsdaten muss sicher und eindeutig sein.

Dies kann u.a. erreicht werden durch:

- a) die Anwendung von internen bzw. nach außen sichtbaren Persistenten Identifikatoren für die digitalen Objekte und ihre Teile, insbesondere die Inhaltsdaten und Metadaten (vgl. 12.1),
- b) Vorhaltung aller zu einem Objekt gehörigen Inhalts- und Metadaten in einem Verzeichnis bzw. einer Datei.

Das Metadatenschema METS bietet die Möglichkeit, ein digitales Objekt, das mit einem Base64-Konverter in eine ASCII-Zeichenfolge konvertiert wurde, innerhalb eines XML-Dokuments mit den Metadaten abzulegen.

C. Infrastruktur und Sicherheit

Infrastruktur und Sicherheit betrachtet die technischen Aspekte des Gesamtsystems sowie die Aspekte der Sicherheit.

13 Die IT-Infrastruktur ist angemessen.

Die IT-Infrastruktur realisiert die Vorgaben des Umgangs mit Objekten in technischer und sicherheitstechnischer Hinsicht. Sie ist für die Gesamtheit aller Objekte zuständig.

13.1 Die IT-Infrastruktur setzt die Forderungen aus dem Umgang mit Objekten um.

Die vom dLZA spezifizierten Anforderungen für den Umgang mit Objekten auf allen Stufen der Verarbeitung werden vom Gesamtsystem implementiert. Dies umfasst die Hauptprozesse (in OAIS: „Funktionalen Entitäten“) Aufnahme (Ingest), Archivablage (Archival Storage) und Nutzung (Access) sowie den Unterstützungsprozess Datenmanagement. Eine Erweiterung kann durch die Ziele des dLZAs erforderlich werden.

Web-Ingest-Modul, Modul für Massen-Ingest im Batch-Betrieb
Speicher-Modul mit Rückgriff auf ein weiteres örtlich ausgelagertes
Speichersystem
Nutzungs-Modul

Ist in der Policy des dLZAs vorgesehen, dass registrierte Nutzer ihre Fotosammlungen, sofern sie als JPEG-Datei vorliegen, selbst in das dLZA spielen dürfen, dann muss das dLZA auch eine entsprechende Upload-Schnittstelle für Nutzer bereit stellen.

Borghoff, Uwe M.: nestor - materialien 3: Vergleich bestehender Archivierungssysteme, 2005. [68]

13.2 Die IT-Infrastruktur setzt die Sicherheitsanforderungen des IT-Sicherheitskonzepts um.

Bei der Umsetzung werden die Sicherheitsanforderungen des Objektmanagements berücksichtigt:
Sicherung der **Integrität** der Objekte, d.h. Schutz vor Modifikationen durch beabsichtigte und unbeabsichtigte menschliche Aktionen sowie technische Unvollkommenheit.
Sicherung der **Authentizität** der Objekte.

Sicherung der **Vertraulichkeit** der Objekte, d.h. Ausschluss unerlaubter Informationsbeschaffung.

Sicherung der **Verfügbarkeit** der Objekte durch die Verfügbarkeit der Funktionalitäten des Objektmanagements (Schutz vor Sabotage, Ausfällen des Systems etc.).

Der Zugang zu geschützten Daten (z.B. Unterlagen des STASI-Untersuchungsausschusses im Archiv) muss durch entsprechende technische Sicherheitsvorkehrungen (z.B. durch Passwörter oder biometrische Zugangsbarrieren) auf die berechtigten Nutzer eingeschränkt sein.

Einsatz qualifizierter digitaler Signaturen nach dem SigG und Zeitstempel für die Archivierung von Patentanmeldungen.

14 Die Infrastruktur gewährleistet den Schutz des digitalen Langzeitarchivs und seiner digitalen Objekte.

Die Infrastruktur schützt die digitalen Objekte vor systembedingten und externen Gefahren. Systembedingte Gefährdungen können sich beispielsweise durch Hardwareprobleme oder den Ausfall einzelner Speichermedien ergeben. Extern muss das dLZA zunächst gegen natürliche Bedrohungen (z.B. Feuer, Wasser, Erdbeben), daneben auch gegen von Menschen verursachte Gefahren geschützt werden. Die Objekte können unmittelbar durch Mitarbeiter oder durch eingeschleuste Schadprogramme (z.B. Viren) gefährdet werden. Zum Schutz der Daten gehört es auch, die nicht beabsichtigte Weitergabe von Informationen durch Programme (Trojaner) oder Menschen (Spionage) zu vermeiden.

Neben den Objekten müssen auch die vom dLZA verwendeten Einrichtungen, Hard- und Software und nicht zuletzt die Mitarbeiter geschützt werden.

Den unterschiedlichen Gefährdungen muss mit einem Paket von technischen (z.B. Virenschutzprogramme) und organisatorischen Maßnahmen (z.B. Zutrittsregelungen) begegnet werden.

Es kann bei einem Brand in einem Kerngebäude der Institution, in der das dLZA betrieben wird, kein Schaden an den Objekten oder Datenverlust entstehen, da es ein entsprechendes Backup an einem weit entfernten Standort gibt, das den Betrieb bei Havarie übernimmt.

BSI: Leitfaden IT-Sicherheit. IT-Grundschutz kompakt, 2007. [69]

BSI: IT-Grundschutz-Kataloge, 2007. [70]

III. Checkliste

Eine absolute Bewertung der Maßnahmen zur Erfüllung der Kriterien ist nicht möglich. Die Bewertung orientiert sich immer an den Zielen des digitalen Langzeitarchivs, wobei die Angemessenheit der Maßnahmen zu prüfen ist.

Es trägt zur Transparenz und damit zur Vertrauensbildung der digitalen Langzeitarchive bei, wenn über die Umsetzung der Kriterien hinaus eine angemessene Dokumentation veröffentlicht wird.

Deshalb sind einerseits die 4 Phasen der Erfüllung (1. Konzeption, 2. Planung und Spezifikation, 3. Umsetzung und Implementation, 4. Evaluierung) zu berücksichtigen, andererseits die Veröffentlichung.

A Organisatorischer Rahmen

1 Das dLZA hat seine Ziele definiert.

1.1 Das dLZA hat Kriterien für die Auswahl seiner digitalen Objekte entwickelt.

1.2 Das dLZA übernimmt die Verantwortung für den dauerhaften Erhalt der durch die digitalen Objekte repräsentierten Informationen.

1.3 Das dLZA hat seine Zielgruppe(n) definiert.

2 Das dLZA ermöglicht seiner(n) Zielgruppe(n) eine angemessene Nutzung der durch die digitalen Objekte repräsentierten Informationen.

2.1 Das dLZA ermöglicht seiner(n) Zielgruppe(n) den Zugang zu den digitalen Objekte.

2.2 Das dLZA stellt die Interpretierbarkeit der digitalen Objekte durch seine Zielgruppe(n) sicher.

3 Gesetzliche und vertragliche Regelungen werden eingehalten.

3.1 Es bestehen rechtliche Regelungen zwischen Produzenten und dem digitalen Langzeitarchiv.

3.2 Das dLZA handelt bei der Archivierung auf der Basis rechtlicher Regelungen.

3.3 Das dLZA handelt bei der Nutzung auf der Basis rechtlicher Regelungen.

4 Die Organisationsform ist für das dLZA angemessen.

4.1 Die Finanzierung des digitalen Langzeitarchivs ist sichergestellt.

4.2 Dem dLZA steht Personal mit angemessener Qualifikation in ausreichendem Umfang zur Verfügung.

4.3 Für das dLZA bestehen angemessene Organisationsstrukturen.

4.4 Das dLZA betreibt eine langfristige Planung.

4.5 Das dLZA reagiert auf substantielle Veränderungen.

4.6 Die Fortführung der festgelegten Aufgaben ist auch über das Bestehen des dLZAs hinaus sichergestellt.

5 Das dLZA führt ein angemessenes Qualitätsmanagement durch.

5.1 Alle Prozesse und Verantwortlichkeiten sind definiert.

5.2 Das dLZA dokumentiert alle seine Elemente nach einem definierten

Verfahren.

B Umgang mit Objekten

6 Das dLZA stellt die Integrität der digitalen Objekte auf allen Stufen der Verarbeitung sicher.

6.1 Aufnahme (Ingest): Das dLZA sichert die Integrität der digitalen Objekte.

6.2 Archivablage (Archival Storage): Das dLZA sichert die Integrität der digitalen Objekte.

6.3 Nutzung (Access): Das dLZA sichert die Integrität der digitalen Objekte.

7 Das dLZA stellt die Authentizität der digitalen Objekte auf allen Stufen der Verarbeitung sicher.

7.1 Aufnahme (Ingest): Das dLZA sichert die Authentizität der digitalen Objekte.

7.2 Archivablage (Archival Storage): Das dLZA sichert die Authentizität der digitalen Objekte.

7.3 Nutzung (Access): Das dLZA sichert die Authentizität der digitalen Objekte.

8 Das dLZA betreibt eine langfristige Planung seiner technischen Langzeiterhaltungsmaßnahmen.

9 Das dLZA übernimmt digitale Objekte von den Produzenten nach definierten Vorgaben.

9.1 Das dLZA spezifiziert seine Übergabepakete (Submission Information Packages, SIPs).

9.2 Das dLZA identifiziert, welche Eigenschaften der digitalen Objekte für den Erhalt von Informationen signifikant sind.

9.3 Das dLZA erhält die technische Kontrolle über die digitalen Objekte, um Langzeitarchivierungsmaßnahmen durchführen zu können.

10 Die Archivablage der digitalen Objekte erfolgt nach definierten Vorgaben.

10.1 Das dLZA definiert seine Archivpakete (Archival Information Packages, AIPs).

10.2 Das dLZA sorgt für eine Transformation der Übergabepakete in Archivpakete.

10.3 Das dLZA gewährleistet die Speicherung und Lesbarkeit der Archivpakete.

10.4 Das dLZA setzt Strategien zum Langzeiterhalt der Archivpakete um.

11 Das dLZA ermöglicht die Nutzung der digitalen Objekte nach definierten Vorgaben.

11.1 Das dLZA definiert seine Nutzungspakete (Dissemination Information Packages, DIPs).

11.2 Das dLZA gewährleistet eine Transformation der Archivpakete in Nutzungspakete.

12 Das Datenmanagement ist dazu geeignet, die notwendigen Funktionalitäten des digitalen Langzeitarchivs zu gewährleisten.

12.1 Das dLZA identifiziert seine Objekte und deren Beziehungen eindeutig und

dauerhaft.

- 12.2 Das dLZA erhebt in ausreichendem Maße Metadaten für eine formale und inhaltliche Beschreibung und Identifizierung der digitalen Objekte.
- 12.3 Das dLZA erhebt in ausreichendem Maße Metadaten zur strukturellen Beschreibung der digitalen Objekte.
- 12.4 Das dLZA erhebt in ausreichendem Maße Metadaten, die alle vom digitalen Langzeitarchiv vorgenommenen Veränderungen an den digitalen Objekten verzeichnen.
- 12.5 Das dLZA erhebt in ausreichendem Maße Metadaten zur technischen Beschreibung der digitalen Objekte.
- 12.6 Das dLZA erhebt in ausreichendem Maße Metadaten, die die entsprechenden Nutzungsrechte und -bedingungen verzeichnen.
- 12.7 Der Erhalt der Paketstruktur ist zu jeder Zeit gegeben.

C. Infrastruktur und Sicherheit

13 Die IT-Infrastruktur ist angemessen.

- 13.1 Die IT-Infrastruktur setzt die Forderungen aus dem Umgang mit Objekten um.
- 13.2 Die IT-Infrastruktur setzt die Sicherheitsanforderungen des IT-Sicherheitskonzepts um.

14 Die Infrastruktur gewährleistet den Schutz des digitalen Langzeitarchivs und seiner digitalen Objekte.

IV. Glossar und Abkürzungen

Archivablage (Archival Storage): eine funktionale Entität nach OAIS, die jene Funktionen und Prozesse beinhaltet, die die Speicherung und Verfügbarkeit der Archivpakete gewährleisten.

Archivpaket (Archival Information Package, AIP): Informationseinheit, bestehend aus inhaltlichen Daten und für die Langzeitarchivierung notwendigen Metadaten, die im dLZA gespeichert wird.

Aufnahme (Ingest): eine funktionale Entität nach OAIS, die jene Funktionen und Prozesse beinhaltet, die die Übernahme der Übergabepakete von den Produzenten bzw. Lieferanten, die Transformation in Archivpakete und die Aufnahme ins Archiv gewährleisten.

Authentizität: das Objekt stellt das dar, was es vorgibt darzustellen.

Daten: Repräsentation von Information in einer formalisierten Art, die die Interpretation, Verarbeitung bzw. den Austausch erlaubt.

Digitales Langzeitarchiv (dLZA): eine Organisation (bestehend aus Personen und technischen Systemen), die die Verantwortung für den Langzeiterhalt und die Langzeitverfügbarkeit digitaler Daten sowie die Bereitstellung für eine bestimmte Zielgruppe übernommen hat. Dabei bedeutet „Langzeit“: über Veränderungen in der Technik (Soft- und Hardware) und in der Zielgruppe hinweg (z.B. für künftige Generationen, auf unbegrenzte Zeit).

Digitales Objekt: logisch abgegrenzte Einheit digitaler Daten. Dies kann ein einfaches Objekt sein, das aus einer einzigen Datei besteht (z.B. ein PDF-Dokument), es kann aber auch ein komplexes Objekt sein, das aus mehreren Dateien besteht (z.B. eine elektronische Zeitschrift mit einzelnen Artikeln als Dateien). Ferner können zu den Daten, die den Inhalt repräsentieren (Inhaltsdaten), weitere Daten hinzukommen (Metadaten), die z.B. der formalen und inhaltlichen Beschreibung, der Strukturbeschreibung, der Interpretierbarkeit oder der Langzeitarchivierung dienen (vgl. Übergabepaket, Archivpaket, Nutzungspaket).

Integrität: 1. die Vollständigkeit der digitalen Objekte, 2. der Ausschluss unbeabsichtigter Modifikationen im Sinne der Erhaltungsregeln. Maßstab für die Integrität sind die als erhaltenswert definierten Eigenschaften eines digitalen Objekts.

Langzeiterhaltungsmaßnahmen (Preservation Planning): die Gesamtheit aller gezielten Verfahren, die angewandt werden, um digitale Objekte dauerhaft zu archivieren und

langfristig verfügbar zu halten. Dazu zählen sowohl Verfahren zum physischen Erhalt der Daten als auch die Anwendung von Migrations- und Emulationsverfahren, die die archivierten Objekte bzw. ihre Umgebungen verändern, um ihre künftige Nutzung zu gewährleisten.

Metadaten: Daten, die Informationen über andere Daten repräsentieren, indem sie z.B. Inhalt, Struktur, Beschaffenheit, Handhabung, Herkunft etc. näher beschreiben. Metadaten können zu unterschiedlichen Zeiten im Lebenszyklus digitaler Objekte entstehen (bei der Produktion, bei der Archivierung, bei der Bereitstellung für die Nutzung...).

Der Begriff wird vor allem im digitalen Umfeld verwendet (z.B. Dublin Core Metadaten), obwohl z.B. Titelaufnahmen in Bibliothekskatalogen, Findbucheinträge etc. ebenfalls als Metadaten aufgefasst werden können. Metadaten werden als Teile der konzeptionellen Einheiten Übergabe-, Archiv-, Nutzungspaket angesehen.

Nutzer: Personen oder Client-Systeme, die mit dem dLZA interagieren, um die durch die digitalen Pakete repräsentierten Information aufzufinden und zu nutzen.

Nutzung (Access): eine funktionale Entität nach OAIS, die jene Funktionen und Prozesse beinhaltet, die die archivierten Informationen den Nutzern zugänglich machen.

Nutzungspaket (Dissemination Information Package, DIP): Informationseinheit, die von einem oder mehreren AIPs abgeleitet wird, und die ein Nutzer als Antwort auf seine Anfrage an das dLZA erhält. Ein Nutzungspaket besteht aus den Daten, die den Inhalt repräsentieren, sowie ggf. den zur Interpretation notwendigen Informationen (z.B. einer Datei im csv-Format und der Beschreibung der Datenstruktur; ein DOS-Programm im Quelltext und eine Emulationssoftware für das Betriebssystem DOS).

OAIS: Referenz-Modell für dLZAe (ISO 14721:2003), das die Kernprozesse eines dLZAs beschreibt (funktionale Entitäten) und ein Informationsmodell zur Verfügung stellt.

Produzent: Personen oder Client-Systeme, die dem dLZA digitale Objekte zum Zweck der Langzeiterhaltung übergeben. Dies sind nicht notwendigerweise die Urheber, sondern können auch Lieferanten der digitalen Objekte sein.

Qualität: Die Qualität eines dLZAs ist der Grad, in dem eine Menge inhärenter kennzeichnender Eigenschaften die spezifizierten Anforderungen erfüllt. Als Anforderungen werden Erfordernisse oder Erwartungen verstanden, die festgelegt, üblicherweise vorausgesetzt oder verpflichtend sind (in Anlehnung an ISO 9000:2000).

Repräsentationsinformation: Information, die notwendig ist, um digitale Daten zu interpretieren (z.B. zu einer Datei das Dateiformat).

Übergabepaket (Submission Information Package, SIP): Informationseinheit, die vom Produzenten an das dLZA abgegeben wird. Die Inhaltsdaten können bereits durch Metadaten angereichert sein.

Verfügbarkeit: Dem Nutzer stehen die Daten zum geforderten Zeitpunkt zur Verfügung.

Vertraulichkeit: Schutz vor unbefugter Preisgabe der Daten.

Vertrauenswürdigkeit (engl. trustworthiness): Eigenschaft eines Systems, gemäß seinen Zielen und Spezifikationen zu operieren (d.h. es tut genau das, was es zu tun vorgibt). Anhand eines Kriterienkatalogs kann die Vertrauenswürdigkeit eines dLZA geprüft und bewertet werden.

Zielgruppe: identifizierbare Gruppe von potentiellen Nutzern mit bestimmten Interessen und Voraussetzungen. Dies kann die Allgemeinheit ebenso sein wie eine Gruppe Fachwissenschaftler. Sie kann heterogen sein und sich aus verschiedenen Nutzergruppen zusammensetzen.

V. Bibliographie

- [1] CCSDS, Consultative Comitee for Space Data Systems: A Reference Model for an Open Archival Information System, Blue Book, 2002,
<http://public.ccsds.org/publications/archive/650x0b1.pdf>, (28.11.2008).
Entwurf zu ISO 14721:2003, International Organization for Standardization: Space data and information transfer systems -- Open archival information system -- Reference model, 2003
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=24683 , (29.11.2008).
- [2] DINI, Deutsche Initiative für Netzwerkinformation e.V., Arbeitsgruppe „Elektronisches Publizieren“ DINI – Zertifikat. Dokumenten- und Publikationsserver 2007, Version 2.0, 2006,
<http://edoc.hu-berlin.de/series/dini-schriften/2006-3/PDF/3.pdf>
(28.11.2008).
- [3] RLG, Working Group on Digital Archive Attributes: Trusted Digital Repositories: Attributes and Responsibilities. An RLG-OCLC Report, Mountain View CA, 2002,
<http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf>,
(28.11.2008).
- [4] RLG – NARA Task Force on Digital Repository Certification: Audit Checklist for Certifying Digital Repositories. Draft for Public Comment, 2005,
<http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070511/viewer/file2433.html>, (28.11.2008).
CRL, OCLC/RLG – NARA Task Force on Digital Repository Certification: Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC), 2007,
<http://www.crl.edu/PDF/trac.pdf>, (28.11.2008).
- [5] DRAMBORA, Digital Repository Audit Method Based on Risk Assessment,
<http://www.repositoryaudit.eu/>, (28.11.2008).
- [6] DCC, Digital Curation Centre,
<http://www.dcc.ac.uk/>, (28.11.2008).
- [7] DPE, Digital Preservation Europe,
<http://www.digitalpreservationeurope.eu/>, (28.11.2008).
- [8] CRL, Centre for Research Libraries: Certification of Digital Archives Project,
<http://www.crl.edu/content.asp?l1=13&l2=58&l3=142>, (28.11.2008).
CRL, Centre for Research Libraries: Long-Lived Digital Collections
<http://www.crl.edu/content.asp?l1=13&l2=58&l3=174>, (02.11.2008).

- [9] DCC, DPE, nestor, CRL: Core Requirements for Digital Archives, Chicago, 2007, <http://www.crl.edu/content/DigArc/metrics2/corerequirements.htm>, (28.11.2008).
- [10] Erpanet–Tagung, Electronic Resource Preservation and Access Network: "Policies for Digital Preservation", 2003, <http://www.erpanet.org/events/2003/paris/index.php>, (28.11.2008).
- [11] Erpanet–Tagung, Electronic Resource Preservation and Access Network: "Appraisal of Scientific Data", 2003, <http://www.erpanet.org/events/2003/lisbon/index.php>, (28.11.2008).
- [12] Wiesenmüller, Heidrun; Jendral, Lars et al.: Auswahlkriterien für das Sammeln von Netzpublikationen im Rahmen des elektronischen Pflichtexemplars: Empfehlungen der Arbeitsgemeinschaft der Regionalbibliotheken. In: Bibliotheksdienst (2004), No 11, S. 1423–1444, http://www.zlb.de/aktivitaeten/bd_neu/heftinhalte/heft9-1204/digitalebib1104.pdf, (28.11.2008).
- [13] BArchG, Bundesarchivgesetz: Gesetz über die Sicherung und Nutzung von Archivgut des Bundes, Koblenz, 2005, <http://www.bundesarchiv.de/benutzung/rechtsgrundlagen/bundesarchivgesetz/index.html>, (28.11.2008).
- [14] Goebel, Jürgen W.; Scheller, Jürgen et al.: nestor – materialien 1: Digitale Langzeitarchivierung und Recht, Frankfurt am Main, 2004, urn:nbn:de:0008-20040916022, http://www.langzeitarchivierung.de/downloads/mat/nestor_mat_01.pdf, (28.11.2008).
- [15] Coyle, Karen: Rights in the PREMIS Data Model. A Report for the Library of Congress, Washington, 2006, <http://www.loc.gov/standards/premis/Rights-in-the-PREMIS-Data-Model.pdf>, (28.11.2008).
- [16] Erpanet–Tagung, Electronic Resource Preservation and Access Network: "Business Models related to Digital Preservation", 2004, <http://www.erpanet.org/events/2004/amsterdam/index.php>, (28.11.2008).
- [17] Digital Longevity Department: Vers van de pers...Kostenmodel digitale bewaring (Kostenmodell für die Langzeitarchivierung), Den Haag, 2006, <http://www.digitaleduurzaamheid.nl/detail.cfm?id=106&sub=nieuws&categorie=0>, (28.11.2008).
- [18] Palm, Jonas: The Digital Black Hole, Stockholm, 2006, http://www.tape-online.net/docs/Palm_Black_Hole.pdf, (28.11.2008).

- [19] Oltmans, Erik; Kol, Nanda: A Comparison between Migration and Emulation in Terms of Costs. In: RLG diginews (2005), Vol 9, No 2, <http://worldcat.org/arcviewer/1/OCC/2007/07/10/0000068902/viewer/file1.html#article0>, (28.11.2008).
- [20] ISO 9000:2005, International Organization for Standardization: Quality management systems. Fundamentals and vocabulary, 2005, http://www.iso.org/iso/catalogue_detail?csnumber=42180, (29.11.2008).
- [21] Liggesmeyer, Peter: Software-Qualität. Testen, Analysieren und Verifizieren von Software, Heidelberg/Berlin, 2002.
- [22] Kneuper, Ralf: CMMI. Verbesserung von Softwareprozessen mit Capability Maturity Model Integration, 2. Auflage, Heidelberg, 2006.
- [23] ITIL: IT – Infrastructure Library, <http://www.itil.org/>, (28.11.2008).
- [24] CCSDS, Consultative Comitee for Space Data Systems: Producer-Archive Interface Methodology – Abstract Standard, Blue Book, Washington, 2004, <http://public.ccsds.org/publications/archive/651x0b1.pdf>, 22.09.2008. Entwurf zu ISO 20652:2006, International Organization for Standardization: Space data and information transfer systems -- Producer-archive interface -- Methodology abstract standard, 2006, http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39577, (29.11.2008).
- [25] Erpanet-Tagung, Electronic Resource Preservation and Access Network: "Workshop on Workflow", 2004, <http://www.erpanet.org/events/2004/budapest/index.php>, (28.11.2008).
- [26] Borghoff, Uwe M.; Rödiger, Peter et al.: Langzeitarchivierung. Methoden zur Erhaltung digitaler Dokumente, Heidelberg, 2003.
- [27] ISO 15489-1:2001, International Organization for Standardization: Information and documentation – Records Management, 2001, http://www.iso.org/iso/catalogue_detail?csnumber=31908, (29.11.2008).
- [28] Shirey, R.: Internet Security Glossary, 2000, www.faqs.org/ftp/rfc/pdf/rfc2828.txt.pdf, (28.11.2008).
- [29] ISO/IEC 15408, International Organization for Standardization: Information technology. Security techniques – Evaluation criteria for IT security, Part 1, 2005; Part 2, 2008, Part 3, 2008, http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40612, (02.11.2008)

- [30] DigiCult, Digital Culture: Integrity and Authenticity of Digital Cultural Heritage Objects, 2002,
http://www.digicult.info/downloads/thematic_issue_1_final.pdf, (28.11.2008).
- [31] Littman, Justin: A Technical Approach and Distributed Model for Validation of Digital Objects. In: D-Lib Magazine (2006), Vol 12, No 5,
<http://www.dlib.org/dlib/may06/littman/05littman.html>, (28.11.2008).
- [32] PREMIS Working Group, Preservation Metadata: Implementation Strategies: Data Dictionary for Preservation Metadata, Version 2.0, Washington, 2008,
<http://www.loc.gov/standards/premis/v2/premis-2-0.pdf>, (28.11.2008).
- [33] Gladney, Henry M.; Bennett, John L.: What Do We Mean by Authentic? What's the Real McCoy? In: D-Lib Magazine (2003), Vol 9, No 7/8,
<http://www.dlib.org/dlib/july03/gladney/07gladney.html>, (28.11.2008).
- [34] InterPARES, The International Research on Permanent Authentic Records in Electronic Systems,
<http://www.interpares.org/>, (28.11.2008).
- [35] nestor – Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen für Deutschland: Digitale Erhaltungsstrategien. In: nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, Version 1.2, 2008, Kapitel 12,
<http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch.pdf>, (28.11.2008).
- [36] Rauch, Carl; Rauber, Andreas: Anwendung der Nutzwertanalyse zur Bewertung von Strategien zur langfristigen Erhaltung digitale Objekte, Wien, 2006,
http://www.ifs.tuwien.ac.at/~andi/publications/pdf/rau_zfbb05.pdf, (28.11.2008).
- [37] Strodl, Stephan; Becker, Christoph et al.: How to Choose a Digital Preservation Strategy: Evaluating a Preservation Planning Procedure, Wien, 2007,
<http://www.ifs.tuwien.ac.at/~strodl/paper/FP060-strodl.pdf>, (28.11.2008).
- [38] DOMEA, Dokumentenmanagement und elektronische Archivierung: Aussonderung und Archivierung elektronischer Akten. Erweiterungsmodul zum DOMEA-Organisationskonzept 2.1, 2005,
http://www.kbst.bund.de/cIn_011/nn_836802/SharedDocs/Anlagen-kbst/Domea/erweiterungsmodul-aussonderung-und-archivierung-elektronischer-akten--pdf,templateId=raw,property=publicationFile.pdf/erweiterungsmodul-aussonderung-und-archivierung-elektronischer-akten--pdf.pdf, (28.11.2008).
- [39] The U.S. National Archives & Records Administration: Disposition of Federal Records. Subpart L -- Transfer of Records to the National Archives of the United

- States, Part 1228, § 1228.270. In: NARA Regulations (2002),
<http://www.archives.gov/about/regulations/part-1228/1.html?template=print>,
(28.11.2008).
- [40] NDAD, National Digital Archive of Datasets: Transfer Procedures (Overview),
2005,
[http://www.ndad.nationalarchives.gov.uk/resources/pdf/xfer_notes_overview.p
df](http://www.ndad.nationalarchives.gov.uk/resources/pdf/xfer_notes_overview.pdf), (28.11.2008).
- [41] DPC, Digital Preservation Coalition: Decision Tree for Selection of Digital
Materials for Long-term Retention, York, 2006,
<http://www.dpconline.org/graphics/handbook/dec-tree.html>, (28.11.2008).
- [42] nestor – Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit
digitaler Ressourcen für Deutschland: Wege ins Archiv: Ein Leitfaden für die
Informationsübernahme in das digitale Langzeitarchiv, erscheint in Kürze.
- [43] Kunze, John: Future-Proofing The Web: What We Can Do Today. In: iPRES –
International Conference on Digital Preservation (2005),
[http://rdd.sub.uni-goettingen.de/conferences/ipres05/download/Future-
Proofing%20The%20Web%20What%20We%20Can%20Do%20Today%20-
%20John%20Kunze.pdf](http://rdd.sub.uni-goettingen.de/conferences/ipres05/download/Future-Proofing%20The%20Web%20What%20We%20Can%20Do%20Today%20-%20John%20Kunze.pdf), (28.11.2008).
- [44] Coy, Wolfgang: nestor – materialien 5: Perspektiven der Langzeitarchivierung
multimedialer Objekte, Frankfurt am Main, 2006,
urn:nbn:de:0008-20051214015,
<http://edoc.hu-berlin.de/series/nestor-materialien/5/PDF/5.pdf>, (28.11.2008).
- [45] Witthaut, Dirk; Zierer, Andrea et al.: nestor – materialien 2: Digitalisierung und
Erhalt von Digitalisaten in deutschen Museen, Frankfurt am Main, 2005,
urn:nbn:de:0008-20041223022,
http://www.langzeitarchivierung.de/downloads/mat/nestor_mat_02.pdf,
(28.11.2008).
- [46] Erpanet-Tagung, Electronic Resource Preservation and Access Network: "File
Formats for Preservation", 2004,
<http://www.erpanet.org/events/2004/vienna/index.php>, (28.11.2008).
- [47] Abrams, Stephen: Digital Formats And Preservation. In: iPRES – International
Conference on Digital Preservation (2005),
[http://rdd.sub.uni-
goettingen.de/conferences/ipres05/download/Digital%20Formats%20And%20Pr
eservation%20-%20Stephen%20Abrams.pdf](http://rdd.sub.uni-goettingen.de/conferences/ipres05/download/Digital%20Formats%20And%20Pr eservation%20-%20Stephen%20Abrams.pdf), (28.11.2008).
- [48] LOC, Library of Congress: Sustainability of Digital Formats. Planning for Library
of Congress Collections, 2006,
<http://www.digitalpreservation.gov/formats/index.shtml>, (28.11.2008).

- [49] ISO 19005-1:2005: Document management – Electronic document file format for long-term preservation, Part 1: Use of PDF 1.4 (PDF/A-1), 2005.
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=38920
- [50] Helfer, Bernward; Lupprian, Karl-Ernst: Dateiformate: Eigenschaften und Eignung für die Archivierung elektronischer Unterlagen. Eine Handreichung für Archivarinnen und Archivare, Wiesbaden und München, 2004,
<http://www.gda.bayern.de/datfor.pdf>, (28.11.2008).
- [51] BSI, Bundesamt für Sicherheit in der Informationstechnik: Auswahl geeigneter Datenformate für die Archivierung von Dokumenten. In: IT-Grundschutz-Kataloge, 9. Ergänzungslieferung, 2007,
<http://www.bsi.de/gshb/deutsch/m/m04170.htm>, (28.11.2008).
- [52] Gutzmann, Ulrike; Kamp, Ulrich et al.: Praktische Lösungsansätze zur Archivierung digitaler Unterlagen: "Langzeitarchivierung" und dauerhafte Sicherung der digitalen Überlieferung. In: Archiv und Wirtschaft (2007), No 1,
<http://www.wirtschaftsarchive.de/akea/handreicherung.htm>, (28.11.2008).
- [53] AK Elektronische Archivierung, VdW Arbeitskreis "Elektronische Archivierung": Matrix zur Bewertung von Dateiformaten, Frankfurt am Main, 2006,
http://www.wirtschaftsarchive.de/akea/Dateiformate_Bewertung_V0.1.xls, (28.11.2008).
- [54] Van Wijk, Caroline; Rog, Judith: Evaluating File Formats for Long-term Preservation. In: iPRES – International Conference on Digital Preservation (2007),
http://rdd.sub.uni-goettingen.de/conferences/ipres07/presentations/Caroline-iPRES2007-11-12oct_CW.pdf, (28.11.2008).
- [55] Steinke, Tobias: Universelles Objektformat: Ein Archiv- und Austauschformat für digitale Objekte, Frankfurt am Main, 2006,
http://kopal.langzeitarchivierung.de/downloads/kopal_Universelles_Objektformat.pdf, (28.11.2008).
- [56] Bischoff, Frank M.: Metadata in preservation: selected papers from an ERPANET seminar at the Archives School Marburg. In: ERPANET Seminar: "Metadata in Digital Preservation" (2004),
<http://nestor.sub.uni-goettingen.de/allg/detail.php?show=764>, (28.11.2008).
- [57] METS, Metadata Encoding and Transmission Standard: METS Schema 1.7 Documentation, 2007,
<http://www.loc.gov/standards/mets/mets-schemadocs.html>, (28.11.2008).
- [58] LMER, Langzeitarchivierungsmetadaten für elektronische Ressourcen, Frankfurt am Main, 2005,
<http://www.d-nb.de/standards/pdf/lmer12.pdf>, (28.11.2008).

- [59] nestor – Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen für Deutschland: Metadatenstandards im Bereich der digitalen Langzeitarchivierung, nationale und internationale Entwicklungen. In: nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, Version 1.2, 2008, Kapitel 10.1, <http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch.pdf>, (28.11.2008).
- [60] NATLIB, National Library of New Zealand: Metadata Standards Framework – Preservation Metadata (Revised), Wellington, 2003, http://www.natlib.govt.nz/files/4initiatives_metaschema.pdf, (28.11.2008).
- [61] DNB, Deutsche Nationalbibliothek, Persistent Identifier: Der Service der Deutschen Nationalbibliothek, <http://www.persistent-identifier.de/>, (28.11.2008).
- [62] PADI, Preserving Access to Digital Information: Persistent Identifiers, <http://www.nla.gov.au/padi/topics/36.html>, (28.11.2008).
- [63] Erpanet-Tagung, Electronic Resource Preservation and Access Network: "Persistent Identifier", 2004, <http://www.erpanet.org/events/2004/cork/index.php>, (28.11.2008).
- [64] nestor – Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen für Deutschland: Persistent Identifier. In: nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, Version 1.2, 2008, Kapitel 13.2, <http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch.pdf>, (28.11.2008).
- [65] Shepherd, Elizabeth; Smith, Charlotte: The Application of ISAD(G) to the Description of Archival Datasets. In: Journal for the Society of Archivists (2000), Vol. 21, No 1, S. 55–86.
- [66] DOMEA, Dokumentenmanagement und elektronische Archivierung : Organisationskonzept, Version 2.1, 2005, http://www.verwaltung-innovativ.de/cln_047/nn_1007684/DE/Organisation/domea__konzept/domea__konzept__node.html?__nnn=true%20%C2%A0, (28.11.2008).
- [67] GDA, Generaldirektion der Staatlichen Archive Bayerns, Metadaten für die Aussonderung und Archivierung digitaler Sachakten München, 2004, <http://www.gda.bayern.de/digpub.htm>, (28.11.2008).
- [68] Borghoff, Uwe M.: nestor – materialien 3: Vergleich bestehender Archivierungssysteme, Frankfurt am Main, 2005, urn:nbn:de:0008-20050117016,

http://www.langzeitarchivierung.de/downloads/mat/nestor_mat_03.pdf,
(28.11.2008).

- [69] BSI, Bundesamt für Sicherheit in der Informationstechnik: Leitfaden IT-Sicherheit. IT-Grundschutz kompakt, Bonn, 2007,
<http://www.bsi.de/gshb/Leitfaden/index.htm>, (28.11.2008).
- [70] BSI, Bundesamt für Sicherheit in der Informationstechnik: IT-Grundschutz-Kataloge, 9. Ergänzungslieferung, 2007,
<http://www.bsi.de/gshb/>, (28.11.2008).

Anhang

Core Requirements for Digital Archives / Kernanforderungen an vertrauenswürdige digitale Langzeitarchive [9]

Das digitale Langzeitarchiv übernimmt die Verantwortung für die dauerhafte Erhaltung und kontinuierliche Pflege der digitalen Objekte für die identifizierten Zielgruppen. Das digitale Langzeitarchiv belegt die organisatorische Beständigkeit (auch in den Bereichen Finanzierung, Personalausstattung, Prozesse), um seine Verantwortung zu erfüllen.

Das digitale Langzeitarchiv verfügt über die erforderlichen Rechte (per Vertrag oder Gesetz), um seine Verantwortung zu erfüllen.

Das digitale Langzeitarchiv besitzt ein effektives und effizientes Geflecht von Grundsätzen (policy).

Das digitale Langzeitarchiv erwirbt und übernimmt digitale Objekte auf der Grundlage definierter Kriterien gemäß seinen Verpflichtungen und Fähigkeiten.

Das digitale Langzeitarchiv stellt die Integrität, Authentizität und Nutzbarkeit der dauerhaft aufbewahrten Objekte sicher.

Das digitale Langzeitarchiv dokumentiert alle Maßnahmen, die während des gesamten Lebenszyklus auf die digitalen Objekte angewendet werden, durch angemessene Metadaten.

Das digitale Langzeitarchiv übernimmt die Bereitstellung der digitalen Objekte.

Das digitale Langzeitarchiv verfolgt eine Strategie zur Planung und Durchführung von Langzeiterhaltungsmaßnahmen.

Das digitale Langzeitarchiv besitzt eine angemessene technische Infrastruktur zur dauerhaften Erhaltung und Sicherung der digitalen Objekte.

Diese gemeinsamen Kriterien stellen die Grundlage weiterer Kooperationen und Kollaborationen dar.